Bioinformatic characterisation of genes associated with coenzyme A biosynthesis in mycoplasmas and expression and isolation of dephospho-coenzyme A kinase from *Mycoplasma* sp. Ms02

by
Tertius Alwyn Ras

Thesis presented in fulfilment of the requirements for the degree of Master of Science (Biochemistry) in the Faculty of Science at Stellenbosch University

Supervisor: Dr A. Botes
Co-Supervisor: Prof. E. Strauss

March 2018

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: March 2018

Abstract

The South African ostrich industry is internationally the leading provider of ostrich products. The increasing popularity of ostrich products has resulted in the adjustment of production strategies, which includes intensifying rearing conditions by using feedlot systems. However, this intensive rearing strategy creates an ideal environment for pathogens, such as mycoplasmas, to spread. There are three *Mycoplasma* species that infect ostriches, which are associated with respiratory diseases. These mycoplasma infections can result in production losses, which not only have an economic impact on the ostrich industry but also significant socio-economic implications. Hence, there is a need for specific and cost-effective treatment against these ostrich-infecting mycoplasmas.

The enzymes involved in the biosynthesis pathway of coenzyme A have long been regarded as potential targets for drug development. These enzymes could, therefore, offer a solution to the control of mycoplasma infections in ostriches. Since the coenzyme A biosynthetic pathway is relatively unexplored in mycoplasmas, the first aim of this study was to determine the presence or absence of enzyme-encoding genes involved in this pathway in *Mycoplasma* species. This was done using a bioinformatics approach. Of the 62 *Mycoplasma* species investigated, there were eight species (13%) found to have none of the enzyme-encoding genes, while the remaining species had at least one. Additionally, twelve enzyme-encoding gene homologues were identified and their predicted identities confirmed by evaluating the conserved and functional motifs and domains. The enzyme-encoding gene found to be most common amongst the investigated species was that of dephospho-coenzyme A kinase (DPCK), the final enzyme in the biosynthesis pathway. Furthermore, there was no correlation between the number of identified coenzyme A biosynthetic pathway enzyme-encoding genes in a species and the phylogeny of the respective proteins. There was also no correlation with the 16S rRNA phylogenetic groupings.

Given the common presence of the DPCK-encoding gene, the second aim of this study was to recombinantly express the DPCK of the ostrich-infecting *Mycoplasma* sp. Ms02 (Ms02) and isolate the protein using a His-tag. The Ms02 DPCK-encoding gene was successfully amplified, cloned and mutated by site-directed mutagenesis to allow for expression in a non-mycoplasma host. However, the soluble expression and isolation of the Ms02 DPCK protein proved to be challenging. Using a variation of methods, the protein was eventually solubilised using a sarkosyl treatment method. A pure isolate of the Ms02 DPCK protein could, however, not be attained when using immobilised metal affinity chromatography

(IMAC) purification. Subsequent activity testing of the isolated DPCK enzyme, using an HPLC-based method, also showed no activity.

Opsomming

Die Suid-Afrikaanse volstruisbedryf is internasionaal die grootste verskaffer van volstruisprodukte. Die toenemende gewildheid van volstruisprodukte het gevolglik veroorsaak dat produksiestrategieë aangepas moet word. Een van hierdie strategieë sluit in die intensifisering van boerderypraktyke deur gebruik te maak van voerkraalstelsels. Hierdie intensiewe boerderystrategieë skep egter 'n ideale omgewing vir patogene, soos mikoplasmas, om in te versprei. Daar is drie *Mycoplasma* spesies wat volstruise infekteer, wat geassosieer word met respiratoriese siektes. Hierdie mikoplasma-infeksies kan verliese in produksie veroorsaak, wat nie net 'n ekonomiese impak op die volstruisbedryf het nie, maar ook beduidende sosio-ekonomiese implikasies het. Daar is dus 'n behoefte aan spesifieke en koste-effektiewe behandeling teen hierdie volstruis-infekterende mikoplasmas.

Die ensieme wat by die biosintetiese padweg van koënsiem A betrokke is, is lank reeds beskou as potensiële teikens vir geneesmiddelontwikkeling. Hierdie ensieme kan dus 'n oplossing bied vir die beheer van mikoplasma-infeksies in volstruise. Aangesien die koënsiem A biosintetiese padweg relatief onverken is in mikoplasmas, was die eerste doelwit van hierdie studie om die teenwoordigheid of afwesigheid van ensiem-enkoderende gene betrokke by hierdie padweg in *Mycoplasma* spesies te bepaal. Dit was gedoen deur gebruik te maak van 'n bioinformatika benadering. Van die 62 *Mycoplasma* spesies wat ondersoek was, was daar agt spesies (13%) wat nie een van die ensiem-enkoderende gene besit het nie, terwyl die oorblywende spesies minstens een gehad het. Daarbenewens is twaalf ensiem-enkoderende geenhomoloë geïdentifiseer en hul voorspelde identiteite bevestig deur die gekonserveerde en funksionele motiewe en domeine te evalueer. Die mees algemene ensiem-enkoderende geen wat gevind was onder die ondersoekde spesies was dié van defosfo-koënsiem A kinase (DPCK), die finale ensiem in die biosintese padweg. Verder was daar geen verband tussen die aantal geïdentifiseerde koënsiem A biosintetiese padweg ensiem-enkoderende gene in 'n spesie en die filogenie van die betrokke proteïene nie. Daar was ook geen korrelasie met die 16S rRNA filogenetiese groeperings nie.

Gegewe die algemene teenwoordigheid van die DPCK geen, was die tweede doelwit van hierdie studie om die DPCK proteïen van die volstruis-infekterende *Mycoplasma* sp. Ms02 (Ms02) uit te druk en die proteïen te isoleer met behulp van ‘n His-merker. Die Ms02 DPCK-enkoderende geen is suksesvol geamplifiseer, gekloon en gemuteer deur posisie-gerigte mutagenese om sodoende ekspressie in ‘n nie-mikoplasma gasheer toe te laat. Die oplosbare uitdrukking en isolasie van die Ms02 DPCK proteïen was egter uitdagend. ‘n Variasie van metodes was gebruik, maar die proteïen was eventueel oplosbaar gemaak met

behulp van 'n sarkosiel behandelingsmetode. 'n Suiwer isolaat van die Ms02 DPCK proteïen kon egter nie verkry word tydens ge-imobiliseerde metaal affiniteits chromatografie (IMAC) suiwering nie. Daaropvolgende aktiwiteitsbepalings van die geïsoleerde DPCK ensiem, met behulp van 'n HPLC-gebaseerde metode, het ook geen aktiwiteit getoon nie.

Acknowledgements

I would like to thank the following people/organisations sincerely for the support throughout this thesis:

Dr A. Botes – It has been a wonderful experience to have you as my supervisor and I have learned so much during this time. Thank you for all the help and advice throughout the duration of my studies. I am also very grateful for your patience and effort in assisting with the writing up of my thesis, especially in the late time of the year.

Prof. E. Strauss – Thank you for the advice and guidance during my studies. I appreciate you welcoming me into your laboratory and providing me with all the necessities to complete my research. I am also very thankful for your input and help with the writing up of my thesis.

Konrad Mostert – My un-biological brother. Thank you for all the help not only in this degree but throughout my entire study journey. I am grateful for all of the insightful conversations, well-deserved (or sometimes not so much) lunch breaks and all-around friendship. I probably would have never made it to this point if it was not for you.

I would like to give a special thanks to my parents, Albie and Linda Ras – I cannot thank them enough for the opportunities they have given me. They have provided me with everything I need to complete my educational journey and I thank them for all the love, support, patience and comfort in times of disappointments.

My grandma, Joan Ras – Thank you for all the support you have given me, financially and emotionally. I am very grateful to have you in my life.

My sister, Chané Ras – Thank you for all the support and late night encouragement. I would also like to thank the rest of my family and friends for all of their support.

My girlfriend, Nina van der Westhuizen – A big thank you for everything! I appreciate all the love and support you have given me throughout this degree and our time together. You were there with me through all of the late nights, weekends and missed out holidays. I am truly blessed to have you by my side. I would also like to thank Nina's parents, Melt and Estelle van der Westhuizen, who accepted me in their home as one of their own children.

Thank you to everyone in the Bellstedt/Botes laboratory, specifically Prof. D.U. Bellstedt, Mrs C. de Villiers, Wiets, Lezan, Dewet and Nic, for all the support, encouragement and friendship. And a special thank you to Dewet for all the assistance with the phylogenetic analyses.

Thank you to all the members of the Strauss laboratory, specifically Dr M. de Villiers, Mrs W.J. Bredell (a.k.a. Helba), Blake, Koketso, Wilmarie and Michelle, for accepting me as one of your own. I will cherish your friendships.

Finally, I am thankful to the National Research Foundation for financial assistance.

Table of Contents

List of Abbreviations

%C	Percentage of cross-linker
%T	Total percentage of acrylamide
aa	Amino acids
acetyl-CoA	Acetyl coenzyme A
ACP	Acyl carrier protein
ADC	Aspartate 1-decarboxylase
ADP	Adenosine diphosphate
ASKHA (superfamily)	Sugar kinase/heat-shock protein 70/actin (superfamily)
ATP	Adenosine triphosphate
BLAST	Basic Local Alignment Search Tool
BLASTP	protein-protein BLAST
bp	base-pairs
BSA	Bovine serum albumin
CAC	Citric acid cycle
CF_3-HMP-PP	4-amino-2-trifluoromethyl 5-hydroxymethylpyrimidine pyrophosphate
CoA	Coenzyme A
CoaBC	Bifunctional phosphopantothenoylcysteine decarboxylase/phosphopantothenoylcysteine synthetase
CoASy	Coenzyme A synthase
contig	Contiguous sequence
CTP	Cytidine triphosphate
DePCoA	Dephospho-coenzyme A
DFP (superfamily)	DNA/pantothenate metabolism flavoprotein (superfamily)
dNTPs	Deoxynucleotide triphosphates
DPCK	Dephospho-coenzyme A kinase
ECF	Energy-coupling factor
GO	Gene ontology
GST	Glutathione-*S*-transferase
HAD	Haloacid dehalogenase
HAD-DPCK	Bifunctional HAD-like protein/dephospho-CoA kinase
HMP-P	4-amino-2-methyl-5-hydroxymethylpyrimidine phosphate
HMP-PP	4-amino-2-methyl-5-hydroxymethylpyrimidine pyrophosphate
HPLC	High-performance liquid chromatography

IMAC	Immobilised metal affinity chromatography
IPTG	Isopropyl-β-D-1-thiogalactopyranoside
kb	kilo base-pairs
KEGG	Kyoto Encyclopedia of Genes and Genomes
KPHMT	Ketopantoate hydroxymethyltransferase
KPR	Ketopantoate reductase
LB	Luria Bertani
MBP	Maltose-binding protein
MEME	Multiple Expectation maximisation for Motif Elicitation
Ms01	*Mycoplasma struthionis* sp. nov.
Ms02	*Mycoplasma* sp. Ms02
Ms03	*Mycoplasma nasistruthionis* sp. nov.
MSA	Multiple Sequence Alignment
msDPCK	Ms02 DPCK-encoding gene
MsDPCK	Ms02 DPCK recombinant protein
MsDPCK_C-His	C-terminal 6xHis-tagged Ms02 DPCK protein
MsDPCK_MBP-His	MBP-6xHis-tagged Ms02 DPCK protein
MsDPCK_N-His	N-terminal 6xHis-tagged Ms02 DPCK protein
NCBI	National Centre for Biotechnology Information
NK (superfamily)	Nucleoside/nucleotide kinase (superfamily)
O/N	Overnight
OD_{600}	Optical density measured at a wavelength of 600 nm
PanK	Pantothenate kinase
PantSH	Pantetheine
PCR	Polymerase chain reaction
PPantSH	4’-Phosphopantetheine
PPAT	Phosphopantetheine adenylyltransferase
PPCDC	Phosphopantothenoylcysteine decarboxylase
PPCS	Phosphopantothenoylcysteine synthetase
PPLO	Pleuropneumonia-like organism
PS	Pantothenate synthase
PSI-BLAST	Position-Specific Iterated BLAST
RAST	Rapid Annotation using Subsystem Technology
RAxML	Randomised Axelerated Maximum Likelihood
RE	Restriction endonuclease
SDM	Site-directed mutagenesis

SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SPR	Subtree-Pruning-and-Regrafting
T_a	Annealing temperature
TB	Terrific Broth
TBST	Tris-buffered saline solution containing Tween
TEMED	N,N,N',N'-Tetramethylethylenediamine
T_m	Estimated annealing temperature

Chapter 1 – Introduction

1.1 Background

Mycoplasmas are cell wall-less bacteria that belong to the class Mollicutes. Members of this class are found throughout nature as parasites in humans, mammals and fish [1]. On a cellular and genomic level, they are among the smallest known organisms capable of self-replication [2]. Their small genome and subsequent lack of various biosynthetic pathways cause them to be dependent on their hosts for nutrients [3].

Three main *Mycoplasma* species have been identified that specifically infect ostriches and are associated with respiratory diseases [4]. These mycoplasma infections can result in production losses due to the cost of treatment, downgrading of carcases, wearing of the ostriches and increased susceptibility to other infections [5]. In some severe cases, death may occur. The cost of feeding is also increased since the ostriches require more time to reach a productive size, although their appetite remains unchanged. All these factors can have a considerable economic impact on ostrich production. In addition, ostrich farming provides job opportunities and is a major contributor to agricultural exports, which have significant socio-economic implications. Consequently, specific and cost-effective treatment against these mycoplasma infections is required.

The enzymes associated with the biosynthesis of coenzyme A (CoA) have long been regarded as potential targets for drug development, albeit mostly against human pathogens [6]. Thus, inhibiting the CoA biosynthetic enzymes could provide a solution to the mycoplasma infections in ostriches. The suitability of CoA biosynthetic enzymes as drug targets can be attributed to the following factors: First, CoA is an essential and ubiquitous cofactor involved in a variety of metabolic processes such as the Krebs-cycle, amino acid biosynthesis, and fatty acid biosynthesis, just to name a few [7, 8]. Second, most organisms cannot import CoA and therefore have to synthesise it anew [8]. Third, key enzymes in the biosynthetic pathway show limited similarity between prokaryotes and eukaryotes with regards to sequence, structure, and mechanism, which allows selective inhibition of the CoA biosynthesis pathway in pathogenic microorganisms [6, 8].

The universal CoA biosynthesis pathway consists of five enzyme-assisted steps, starting with pantothenate as substrate and ending with CoA as a product (Figure 1.1). The enzymes in sequential order are pantothenate kinase (PanK), phosphopantothenoylcysteine synthetase (PPCS), phosphopantothenoylcysteine decarboxylase (PPCDC), phospho-pantetheine adenylyltransferase (PPAT) and dephospho-coenzyme A kinase (DPCK) [7].

However, a salvage pathway has been identified where PanK uses pantetheine as a substrate and bypasses PPCS and PPCDC to provide a substrate for PPAT [7]. This might be the case in some *Mycoplasma* species, in which genes encoding for PPCS and PPCDC have not yet been identified [9].

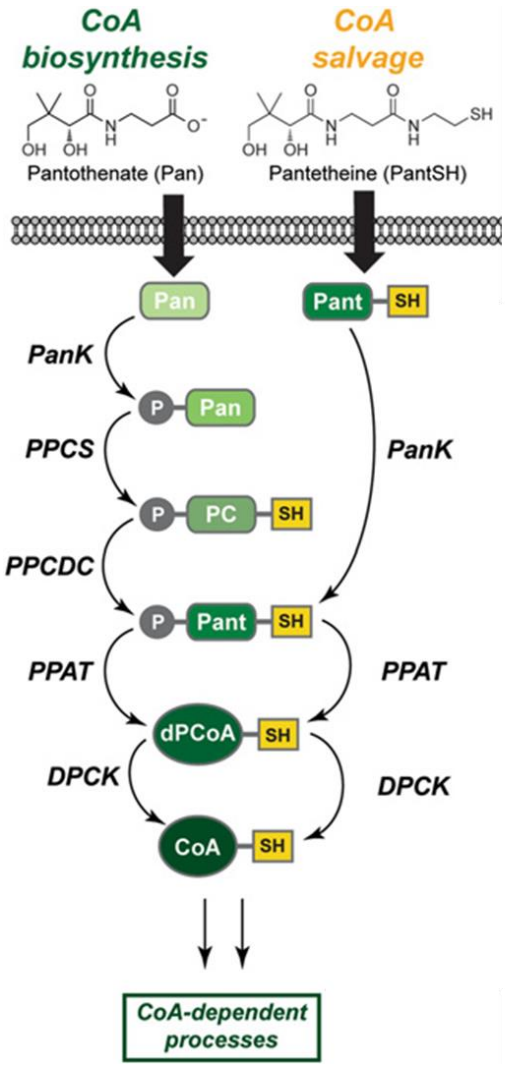


Figure 1.1 A simplification of the universal five-step CoA biosynthesis pathway and the salvage pathway. In the universal pathway, the conversion of pantothenate to CoA is catalysed by PanK, PPCS, PPCDC, PPAT and DPCK. Conversely, in the salvage pathway only three of the five enzymes, namely PanK, PPAT and DPCK, catalyses the conversion of pantetheine to CoA. Adapted from Moolman *et al*. [6].

The number of annotated genes that code for enzymes involved in the CoA biosynthesis pathway is inconsistent amongst mycoplasma genomes. If it is assumed that a missing gene is due to the enzyme not being required it would imply that, depending on the *Mycoplasma* species, not all of the enzymes in the mycoplasma CoA biosynthesis pathway would be a suitable drug target. However, missing genes could be due to insufficient annotation, which is a notable problem within mycoplasma genomes.

1.2 Aims and objectives

The CoA biosynthesis pathway in mycoplasmas is relatively unexplored. Hence, the first aim of this study was to determine the presence or absence of enzyme-encoding genes involved in the CoA biosynthesis pathway in *Mycoplasma* species using a bioinformatics approach. This would allow the identification of enzymes in the pathway that would be suitable as potential drug targets. To this end, the following objectives were set:

- Identify the CoA biosynthesis enzyme-encoding genes currently annotated in the mycoplasma genomes.
- Confirm the presence/absence of the enzyme-encoding genes not annotated in the genomes.
- Determine the conserved and functional motifs and domains in newly identified genes, to confirm their identity.
- Determine the phylogenetic relationship of these *Mycoplasma* species and their respective CoA biosynthetic enzymes.

Given that the most conserved enzyme in the pathway within mycoplasmas (based on currently annotated genomes) seems to be DPCK, the second aim of this study was to recombinantly express and isolate the DPCK of an ostrich-infecting mycoplasma to allow future characterisation and comparison to previously characterised DPCK's of other organisms. For this purpose the DPCK-encoding gene of *Mycoplasma* sp. Ms02 was chosen and the following objectives were set accordingly:

- Amplify the DPCK-encoding gene of *M.* sp. Ms02.
- Clone the amplified gene into a suitable expression vector.
- Evaluate the expression and isolation of the *M.* sp. Ms02 DPCK protein.
- Perform preliminary tests for activity of the *M.* sp. Ms02 DPCK protein.

1.3 Thesis layout

Chapter 2 of this thesis consists of a literature review, in which mycoplasmas and the CoA biosynthesis pathway are discussed. This includes the history and characteristics of mycoplasmas, as well as the existing treatment against avian mycoplasmas, followed by an overview of CoA, the biosynthesis thereof and its potential as a drug target.

The results of the bioinformatics analysis of the CoA biosynthesis pathway in *Mycoplasma* species are presented in Chapter 3, followed by the results of the expression and isolation of the *M.* sp. Ms02 DPCK protein in Chapter 4.

Then, in Chapter 5, conclusions and future perspectives of this study are given, followed by the list of references. The thesis is finalised with Appendices containing the supplementary data of Chapter 3 (Appendix 1) and Chapter 4 (Appendix 2).

Chapter 2 – Literature Review

This chapter provides an overview of mycoplasmas and the coenzyme A (CoA) biosynthesis pathway. First, a brief history of mycoplasmas is given, followed by the characteristics and pathogenicity of these organisms. Moreover, ostrich-infecting mycoplasmas and the current treatment for these infections are discussed. Next, the importance of CoA is highlighted and the biosynthesis thereof is explained. Finally, the recognition of CoA biosynthesis as a potential drug target is discussed.

2.1 Mycoplasmas

2.1.1 History

Louis Pasteur was first to identify mycoplasmas as a microbial entity, which was the etiological agent for contagious bovine pleuropneumonia [10]. However, he could never grow it in a nutrient broth nor observe it microscopically. Then in 1898, the first cultivation of a mycoplasma was reported [11]. This mycoplasma was known as a pleuropneumonia-like organism (PPLO), which was later classified as *Mycoplasma mycoides*. The morphologic properties of this PPLO were described in 1910, followed by filtration studies in 1929 [12]. Their small size and ability to pass through filters, which prevented bacteria from passing through, led to the mycoplasmas being regarded as viruses for years [13].

In 1941, mycoplasmas (at this time still referred to as PPLOs), following the isolation from sewage, humans and other animals, were recognised as a novel group of parasitic and saprophytic microorganisms, which possesses unusual properties that distinguish them from viruses, ordinary bacteria, and rickettsiae [10, 12]. Then in 1956 Edward and Freundt [14] proposed the classification and nomenclature of mycoplasmas, which is the classification scheme we use today.

However, the 1950s and 1960s were filled with literature opposing or supporting the description of mycoplasmas as bacterial L-forms (bacteria that have lost their cell wall partially or entirely) given that they are both cell-wall-less bacteria with an odd “fried-egg” colony shape [13]. This dispute finally came to an end in the late 1960s with genomic analysis data using DNA-hybridization, which disproves the relationship between stable L-forms of present-day walled bacteria and mycoplasmas [15, 16]. Nonetheless, from a long-range evolutionary viewpoint, associating mycoplasmas and L-forms are not entirely incorrect [13].

Our progress of knowledge regarding the ultrastructure, genome, metabolic pathways and cell membrane of mycoplasmas in the 1960s and 1970s resulted in the recognition of mycoplasmas as the smallest organisms capable of self-replication [13]. This sparked up a debate on the position of mycoplasmas in the evolutionary scheme, with some proposing that they should be at the root of the phylogenetic tree [17] and others suggesting that they originated from walled-bacteria via degenerative evolution [18]. The latter was confirmed in 1980 based on the phylogenetic analysis of 16S rRNA genes by Woese *et al.* [19]. Then 15 years later, the first complete sequencing and annotation of a mycoplasma genome (*Mycoplasma genitalium*) was published [20].

2.1.2 Cell morphology

Mycoplasmas possess unique properties such as resistance to penicillin, a “fried-egg” shape colony structure, and sensitivity to detergents and osmotic shock, which can be attributed to their lack of a cell wall [13]. Thin sections on mycoplasmas have shown that the cells fundamentally consist of three organelles, namely the cell membrane, ribosomes and a densely packed, circular double-stranded DNA molecule [13, 21].

The cell sizes of mycoplasmas are within the range of 200-700 nm, with a predominantly spherical cell shape [22]. However, there are many mollicutes that exhibit a variety of cell shapes, for instance, flask-shaped cells with terminal tip structures, pear-shaped cells, filaments of varying lengths (with some branching), and even helical filaments [13]. Sustaining these structures without a rigid cell wall indicates the presence of a cytoskeleton in mycoplasmas [21]. In fact, multiple mycoplasmas were treated with detergents, and a network of filamentous rods and threads was observed [13]. These cytoskeleton-like structures are believed to play a role in regulating cell shape, along with participating in cell division [23].

2.1.3 Taxonomy and Phylogeny

The genus *Mycoplasma* belongs to the class Mollicutes, which name is derived from the Latin words “*mollis*” (meaning ‘soft’) and “*cutis*” (meaning ‘skin’) [24]. ‘Mycoplasmas’ or ‘mollicutes’ are trivial terms that have been used interchangeably for referring to any species in the Mollicutes class [13]. The Mollicutes class consists of four orders, six families and ten genera (Table 2.1) [2, 25, 26].

The comparison of the 16S rRNA sequences of mycoplasmas by Woese *et al*. [19] provided the basis of mycoplasma phylogeny. There are other conserved genes, such as the *tuf* gene,

that have also been used to set up phylogenetic trees of mycoplasmas, but the 16S rRNA sequences are still preferred [27].

Based on the 16S rRNA sequences, the mycoplasmas originated from a node in a branch of low G+C containing gram-positive bacteria via degenerative evolution. This particular branch contains species from *Lactobacillus, Streptococcus,* and *Bacillus*, as well as two *Clostridium* species [19, 28]. Maniloff [28] suggested an evolutionary scheme in which mycoplasmas arose from the *Streptococcus* phylogenetic branch approximately 600 million years ago. This can be considered as a fairly recent development on the evolutionary scale [13].

Table 2.1 The classification of genera in the class Mollicutes

Order	Family	Genus
Mycoplasmatales	*Mycoplasmataceae*	*Mycoplasma* *Ureaplasma*
Entomoplasmatales	*Spiroplasmataceae* *Entomoplasmataceae*	*Spiroplasma* *Entomoplasma* *Mesoplasma*
Acholeplasmatales	*Acholeplasmataceae* *Incertae sedis*	*Acholeplasma* *Candidatus* Phytoplasma
Anaeroplasmatales	*Anaeroplasmataceae*	*Anaeroplasma* *Asteroleplasma*

Based on the 16S rRNA sequences, the mollicutes are split into five phylogenetic groups, known as the hominis-, pneumoniae-, spiroplasma-, anaeroplasma-, and asteroleplasma groups (Figure 2.1) [25, 29, 30]. The *Mycoplasma* species are distributed amongst the former three groups [29].

2.1.4 Genome

Mycoplasmas are thought to have lost many non-essential genes, including those associated with cell wall synthesis, by means of degenerative evolution [24]. Consequently, their genomes are the smallest among all prokaryotes [31]. Data have shown genome sizes for mollicutes ranging from less than 600 kb to over 2200 kb; compared to 4720 kb for *Escherichia coli* [21, 24]. In fact, genome sizes of *Mycoplasma* species range from 580-1350 kb, while *Spiroplasma* species range from 780-2220 kb [21].

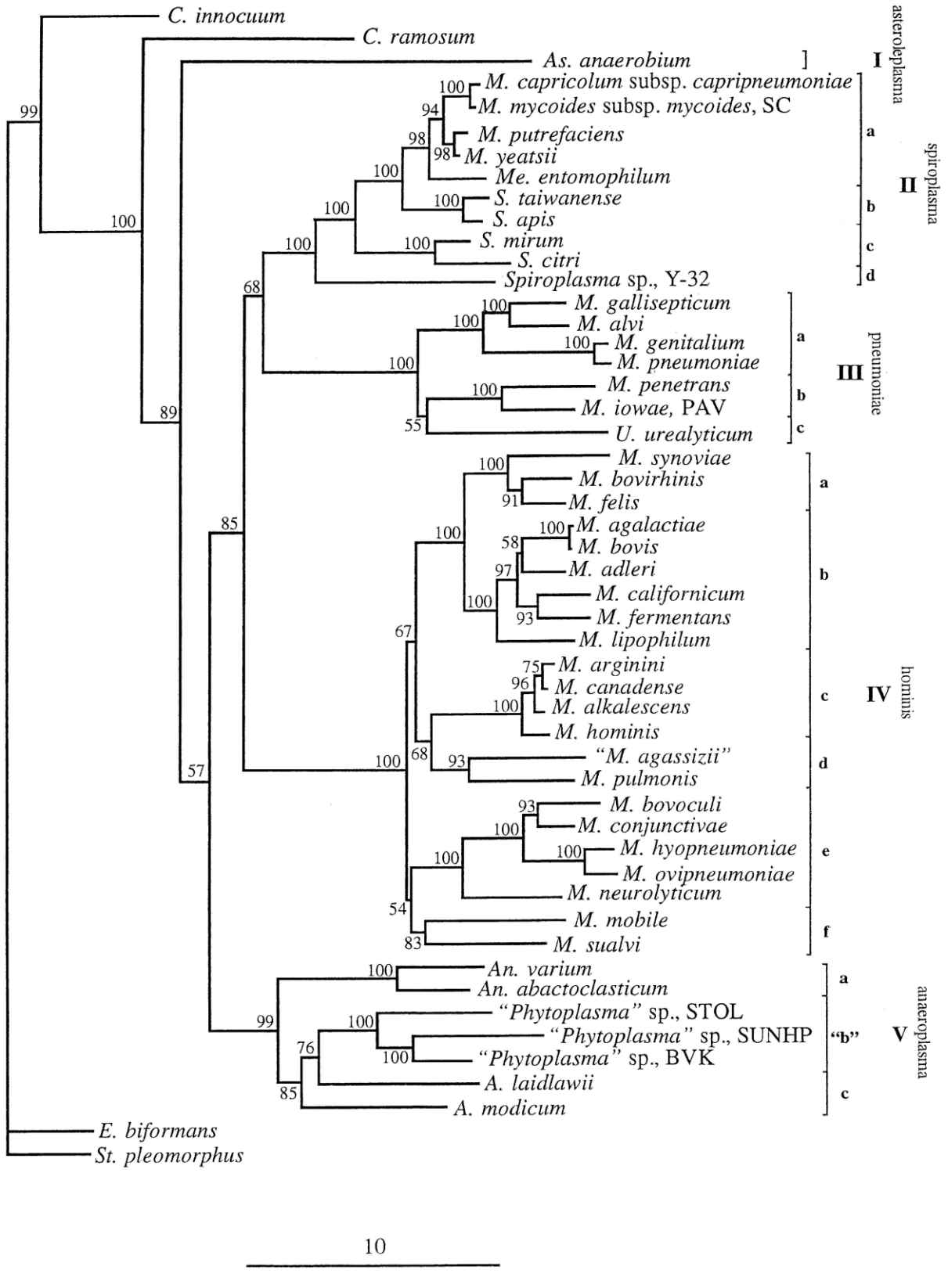


Figure 2.1 A phylogenetic tree of mycoplasmas based on 16S rRNA sequences illustrating the five groups. Representatives of closely related genera *Clostridium* and *Eubacterium* are also included in the tree. Obtained from Johansson *et al*. [30].

In addition to having a small genome, mycoplasmas also have a characteristically low G+C content in their genomes, which ranges between 23 mol% and 40 mol% [32]. However, the G+C content is unevenly distributed along the genome [21, 33]. For instance, in *M. genitalium* the average G+C content of the genome is 32 mol%, yet the G+C content for its tRNA genes is 52 mol%, and 44 mol% for its rRNA genes [20, 34].

Mycoplasmas also make use of a different genetic code, compared to other bacteria. The UGA codon, which is universally used as a stop codon, is used to encode tryptophan in mycoplasmas [22, 33]. Experimentally, this can generate problems, particularly for heterologous expression of cloned mycoplasma genes, which will lead to truncated gene products.

2.1.5 Metabolic pathways

The restricted coding capacity of mycoplasmas, caused by their small genome, has limited their metabolic capabilities. Metabolic activities that have been identified seem to be related to energy production, rather than supplying substrates for biosynthetic pathways [13]. Thus, mycoplasmas are unequipped to synthesise molecules such as fatty acids, some amino acids, cholesterol, purines, and pyrimidines, resulting in their parasitic lifestyle [35].

Furthermore, they lack a complete citric acid cycle (CAC) and possess no cytochromes and quinones, which rules out the use of the highly effective oxidative phosphorylation pathway for generating ATP [36]. Hence, the identified energy-yielding pathways of mycoplasmas produce low ATP yields with relatively large quantities of metabolic end-products. In some cases, the mycoplasma may deplete the host tissue of substrates required for these energy-yielding pathways, which can be a contributing factor in the pathogenicity of these mycoplasmas [21].

Mycoplasmas are divided into fermentative- and nonfermentative organisms, based on their ability to metabolise carbohydrates [2]. Members of the fermentative mycoplasmas use glycolysis for ATP formation, whereas most nonfermentative, and some fermentative, mycoplasmas make use of the arginine dihydrolase pathway [2, 21]. However, the use of the arginine dihydrolase pathway as a sole energy-producing source in nonfermentative mycoplasmas has often been questioned [34, 37].

Another possible energy-generating mechanism in mycoplasmas is based on the production of ATP from adenosine 5’-diphosphate and acetyl phosphate by acetate kinase, coupled with the production of acetyl phosphate from acetyl coenzyme A (acetyl-CoA) by phosphate acetyltransferase [21]. Both of these enzymes are ordinarily found in fermentative and

nonfermentative mycoplasmas. Acetyl-CoA can be produced in mycoplasmas via oxidative decarboxylation of pyruvate [37]. Nevertheless, the contribution of this pathway in energy metabolism of mycoplasmas has yet to be critically evaluated.

2.1.6 Host distribution and habitat

The absence of various biosynthetic pathways in mycoplasmas, render them dependent on their hosts for nutrients [32]. They are found extensively in nature as parasites of humans [38], mammals and birds [24, 39], reptiles [40], fish, arthropods and plants [41, 42]. The popular habitats of mycoplasmas in humans and animals are the mucous surfaces of urogenital- and respiratory tracts, the eyes, mammary glands, alimentary canal, and the joints [21]. The spiroplasmas and phytoplasmas mostly reside in the gut, salivary glands, and hemocoel of arthropods, where they may also be introduced into the phloem tissues of plants by sap-sucking insects [41].

Mycoplasmas usually exhibit high host and tissue specificity, supporting their parasitic or saprophytic lifestyle and particular growth requirements [21]. However, there are several cases of mycoplasmas being present in unfamiliar hosts or tissues that differ from their normal habitats [42].

2.1.7 Pathogenicity

Many mollicutes are commensal organisms, and in some arthropods, they may even be considered as symbionts [43]. Conversely, pathogenic mycoplasma infections usually follow a chronic course and are rarely of the fulminant type [13]. The focus from here on out will be on animal mycoplasmas, where the term ‘mycoplasma’ will be restricted to the genus *Mycoplasma*.

Pathogenic mycoplasmas are connected to many diseases in farm animals. In ruminants, such as cattle, goats and sheep they are known to cause contagious pleuropneumonia, mastitis, conjunctivitis, and even arthritis [39]. Mycoplasma pathogens are also linked to pneumonia and arthritis in swine, as well as pleuritis in horses. Furthermore, in poultry, they are associated with chronic respiratory disease, sinusitis and air sacculitis [39].

2.1.8 Ostrich-infecting mycoplasmas

Botes *et al.* [5] identified three ostrich-infecting *Mycoplasma* species, which are associated with respiratory diseases. Because mycoplasmas are typically named in relation to their host or disease pathology, the three species were temporarily named, Ms01, Ms02 and Ms03 (Ms = ‘*Mycoplasma struthiolus*’), after their host, *Struthio camelus*. Ms01 and Ms03 have

since been described as *Mycoplasma struthionis* sp. nov. and *Mycoplasma nasistruthionis* sp. nov., respectively [44]. However, these names have not been formally accepted. Accordingly, the three species will be referred to as Ms01, Ms02 and Ms03.

A phylogeny based on 16S rRNA sequences (Figure 2.2), indicated that these three *Mycoplasma* species fall in two different clades, in which Ms02 and Ms03 are more closely related to each other, compared to their relation to Ms01 [5, 45]. Nevertheless, all three species are located in the hominis group.

Infections of these mycoplasmas in ostriches are respiratory by nature and can lead to reduced production, due to the wearing of the ostriches, downgrading of carcases, cost of treatment, and increased feeding costs because of slow growth rates [4, 5, 46]. Furthermore, mycoplasma infections increase susceptibility to secondary infections, which can result in mortalities.

2.1.9 Existing treatment of avian mycoplasmas

Treatment of mycoplasma infections has remained a recurring issue due to the flexibility of these organisms. For the purpose of this study, the focus will be on the current treatment of avian mycoplasma infections, specifically antibiotics and vaccines.

2.1.9.1 Antibiotics

Antibiotic treatment has been used as an effective tool against avian mycoplasmas for the reduction of clinical signs, lesions and egg transmission [47]. The antibiotics that target cell wall synthesis, for example, penicillin and cephalosporin, have no effect on mycoplasmas due to their non-existent cell wall. They are, however, sensitive to most broad-spectrum antibiotics, such as tetracyclines (chlortetracycline, oxytetracycline and doxycycline), quinolones (imequil, enrofloxacin, norfloxacin and danofloxacin), macrolides (tylosin, erythromycin, kitasamycin, spiramycin and lincomycin) and tiamulin [48]. In ostriches, tylosin, doxycycline, oxytetracycline, chlortetracycline and lincomycin have been used successfully against mycoplasma infections [49].

Nevertheless, current antibiotic treatment only manage the mycoplasma infections and not completely eradicate them from the flock [50, 51]. Although the use of broad-spectrum antibiotics as a treatment against avian mycoplasma infections can be effective and useful for preventing economic losses, it should not be considered as a long-term solution [47]. Therefore, specific treatment against avian mycoplasmas is required.

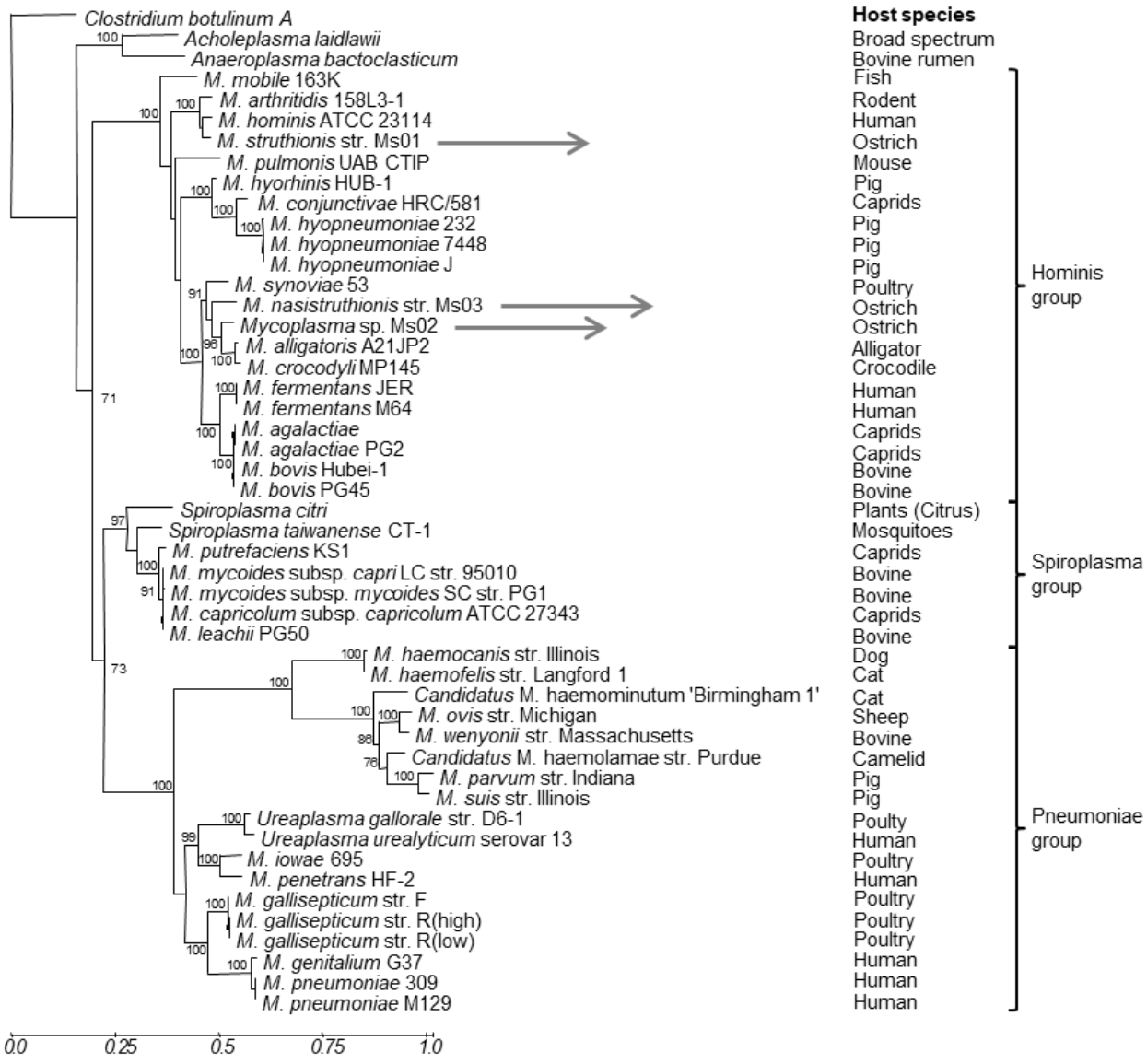


Figure 2.2 A phylogenetic tree of different representative mycoplasmas in three groups (hominis, spiroplasma and pneumoniae) based on 16S rRNA sequences. The three ostrich-specific mycoplasmas have been emphasised with arrows. In addition, the column on the right indicates the host of the respective species. Adapted from Wium *et al*. [45].

2.1.9.2 Vaccines

There are several developed vaccines for mycoplasma infections in poultry, specifically against *Mycoplasma gallisepticum* and *Mycoplasma synoviae*. The current commercially available vaccines used for these infections are either live attenuated vaccines or inactivated vaccines (bacterins) with an oil emulsion adjuvant [47, 52]. A recombinant Vectormune FP MG vaccine, which is a genetically modified fowl pox viral vaccine that expresses *M. gallisepticum* antigens, has also been developed [53] and is available in South Africa. However, this vaccine does not provide sufficient protection when used on its own [54] and should, therefore, be used in combination with a live attenuated vaccine.

Currently, there are no commercially available vaccines against ostrich-infecting mycoplasmas. The oil-emulsified bacterins often cause tissue lesions [55], rendering it unfeasible for the use in ostriches. However, in 2009, Pretorius tested the hypothesis that *M. gallisepticum* and *M. synoviae* poultry vaccines, specifically two bacterins removed from their oil-emulsion, could provide protection against ostrich-specific mycoplasma infections [56]. This hypothesis was proven to be false, as the administered vaccines were reported to offer no protection against the ostrich-specific mycoplasmas due to the absence of antibody cross-reactivity.

The development of a DNA vaccine has been the main focus for a solution against ostrich-infecting mycoplasmas in recent years. The *OppA* gene was subsequently identified as a candidate gene for use in a DNA vaccine against Ms01 [56], Ms02 [57], and Ms03 [45, 58]. This gene codes for the surface-exposed substrate binding domain of the oligopeptide permease transport system. All three of the identified genes were each cloned into commonly used DNA vaccine vectors, namely VR1020, VR1012, and pCI-neo [56, 58, 59].

The constructed DNA vaccines were subsequently used in vaccination trials to determine their ability to elicit an anti-OppA immune response. Earlier vaccination trials against Ms01 and Ms03 were inconclusive due to unforeseen avian influenza infections, stress-induced handling of ostriches and possible ineffective doses [58, 60, 61]. Then, later vaccination trials against Ms01 [49] and Ms03 [62], demonstrated the ability of the VR1020_*OppA* DNA vaccine construct to elicit an anti-OppA immune response in ostriches. Thus far, no vaccination trials against Ms02 have been performed.

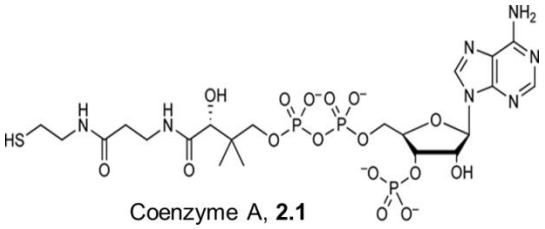
Although the vaccines were able to elicit an immune response, their ability to provide protection against ostrich-infecting mycoplasmas has yet to be confirmed. Thus, as an alternative, prominent metabolic pathways could be evaluated to identify novel proteins as possible drug/vaccine targets. Recently, the biosynthesis pathway of coenzyme A has been identified as a potential target for drug and vaccine development in two *Mycoplasma* species [63, 64].

2.2 Coenzyme A

2.2.1 Coenzyme A and its relevance

Coenzyme A (CoA, **2.1**) was first described as a coenzyme functioning as an acyl carrier in 1947 [65]. Recently, it has been described as an essential omnipresent cofactor involved in numerous metabolic reactions in many, if not all, living organisms [8, 66]. CoA (and the phosphopantetheine tether derived from CoA, which is present on an array of carrier

proteins) serves as the primary carrier of metabolites with carboxylic acid groups [7]. The most common compounds carried include long and short carbon chain acids, amino acids, and Krebs cycle metabolites. Even so, most CoA molecules are associated with the transfer of acetyl groups amongst various small- and macromolecules at any given time [7]. Acetyl-CoA occupies a vital role in multiple cellular processes: as a precursor of anabolic reactions, as a metabolic intermediate, as a key determinant of protein acetylation, and as an allosteric regulator of enzymatic activities [67, 68].



Coenzyme A, **2.1**

The main biological functions of CoA are executed by the ability of its terminal thiol group to form thioester linkages with carboxylic acids [69]. This can be attributed to the distinctive reactivity and chemistry of thioesters and thiols in comparison to other common biologically applicable functional groups [7]. In a biological context, the thiolate anion (a very good nucleophile) is easily formed due to the increased acidity of thiols (pK_a ~9-10), which assists with the acylation reactions where CoA itself is involved [7]. Thioesters, however, can react as electrophiles toward attack by heteroatom nucleophiles (like amines and alcohols), as well as carbon nucleophiles (like enolates). Therefore, favourable thermodynamic and kinetic driving forces allow acyl transfer form CoA thioesters to these groups to occur readily. Despite its reactivity profile, thioesters are no more susceptible to hydrolysis than their oxygen ester equivalents at neutral pH, making them (and therefore also CoA-based thioesters) the perfect form in which acyl groups can be carried in aqueous surroundings [70, 71]. Furthermore, the α-carbon of CoA thioester derivatives can react as a nucleophile upon deprotonation [69]. This dual mode reactivity (Figure 2.3) is further motivation for the vast utility of CoA as a biological cofactor.

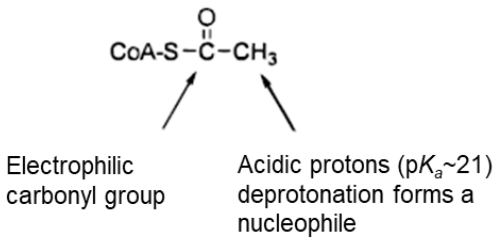


Figure 2.3 The two common modes in the reactivity of acetyl-CoA. Adapted from Mishra *et al.* [69].

2.2.2 Coenzyme A biosynthesis

There are two main substrates that can be used to synthesise CoA, namely pantothenate and pantetheine (PantSH). The former is the common precursor (in most organisms) for the universal five-step biosynthesis pathway, which was first formulated by Brown *et al.* [72]. Pantothenate, also known as pantothenic acid or vitamin B_5, can either be taken up from the extracellular environment or synthesised *de novo* [73]. PantSH can also be obtained from the environment, in addition to it being a product of CoA dissimilation [74]. Yet, the use of PantSH as a substrate for CoA production will require the application of a truncated pathway, which uses only three of the enzymes in the five-step pathway. This is referred to as the CoA salvage pathway [75].

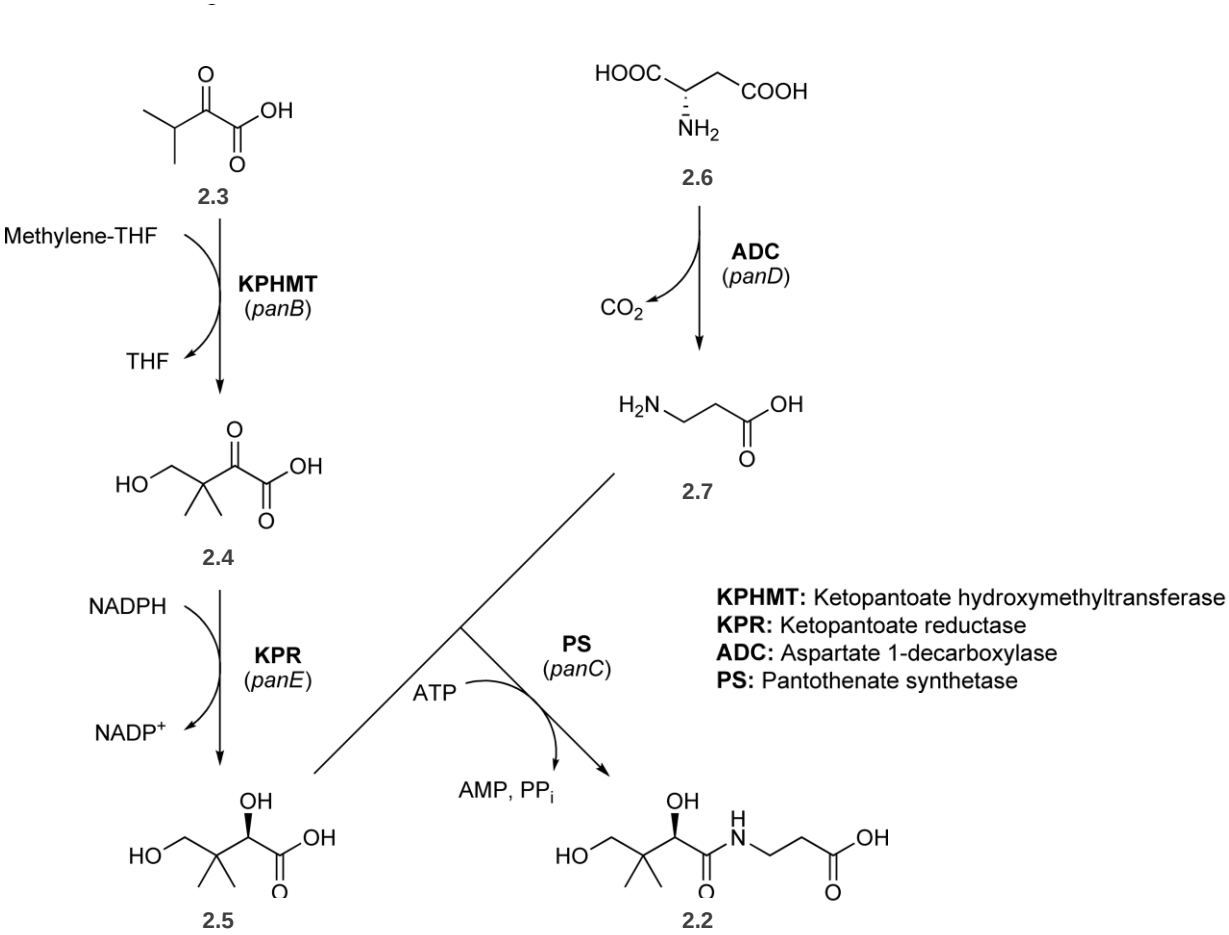
2.2.2.1 Pantothenate biosynthesis

Scheme 2.1 is an illustration of the general pantothenate biosynthetic pathway in *E. coli*. All plants, fungi, and most bacteria have the ability to biosynthesise pantothenate (**2.2**, Scheme 2.1) themselves, which is often overproduced [66]. As a result, pantothenate is found nearly everywhere in biology. *E. coli*, for example, produces and secretes 15 times more pantothenate than it requires to meet its intracellular CoA biosynthetic needs [76]. Thus, the general surplus of pantothenate in the environment can sustain the needs of other organisms that are unable to biosynthesise pantothenate.

Maas *et al.* [77, 78] showed that the formation of pantothenate is an ATP-dependent coupling reaction with pantoic acid (**2.5**) and β-alanine (**2.7**), with AMP and pyrophosphate as side products. Pantoic acid is formed by a two-step enzyme assisted pathway. First, ketopantoate hydroxymethyltransferase (KPHMT) transforms ketoisovaleric acid (**2.3**) into ketopantoic acid (**2.4**) with the help of the N^8,N^{10}-methylene tetrahydrofolate as a cofactor, which provides the hydroxymethyl group. The NADPH-dependent ketopantoate reductase (KPR) then reduces ketopantoic acid to pantoic acid. The source of β-alanine, however, varies according to the organism. For instance, in *E. coli* the only means of producing this amino acid is by decarboxylation of aspartic acid (**2.6**), which is catalysed by the pyruvoyl-dependent aspartate 1-decarboxylase (ADC) [79]. Conversely, in *Saccharomyces cerevisiae,* it can be acquired via the degradation of spermidine and spermine, whereas other organisms may obtain it via degradation of uracil or cysteine, or an alternative pathway of propionate catabolism [80]. Finally, pantothenate synthase (PS) catalyses the coupling of pantoic acid with β-alanine to produce pantothenate.

Because PS catalyses the final step of pantothenate biosynthesis, it has recently received attention as a target for inhibition in organisms dependent on the pathway, where

environmental supply is limited [81, 82]. However, this biosynthetic pathway is not a potential target in the scope of this study. This is due to the lack of various metabolic pathways in mycoplasmas, accompanied by the general availability of pantothenate in the surrounding environment.



Scheme 2.1 The general biosynthetic pathway of pantothenate in *E. coli*, with abbreviated enzyme names in bold and the corresponding enzyme-encoding genes in italics. Ketoisovaleric acid (**2.3**); Ketopantoic acid (**2.4**); Pantoic acid (**2.5**); Aspartic acid (**2.6**); β-alanine (**2.7**); Pantothenate (**2.2**). Adapted from Strauss [7].

2.2.2.2 Transport of Pantothenate (and Pantetheine)

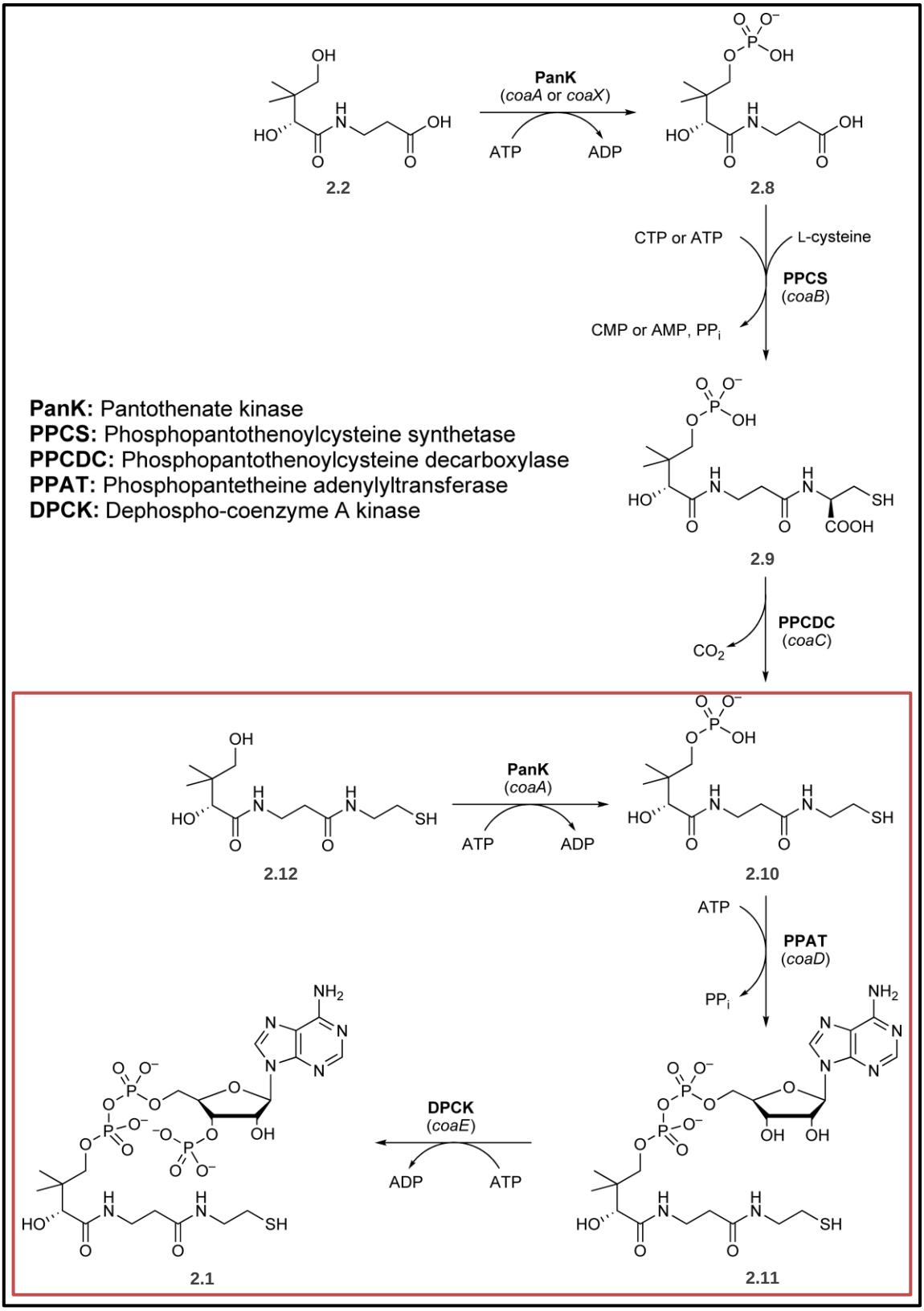
Most organisms, regardless of their ability to synthesise it from anew, possess a transport system that actively facilitates the import of pantothenate [83]. The best-characterised transport system is pantothenate permease (*panF* gene product) in *E. coli* [84, 85]. This system consists of a unidirectional sodium-based symporter specific for pantothenate, which has a maximum velocity of 1.6 pmol/min per 10^8 cells [84, 86]. The uptake of pantothenate is coupled to the symport of sodium with a K_m value for sodium of 0.8 mM and for pantothenate of 0.6 µM [84]. Overexpression of this permease results in a 10-fold rate of transport increase, along with an intracellular pantothenate concentration increase [85]. However, the

transport rate is unaffected by the intracellular CoA concentration, nor are the CoA levels affected by the increase of intracellular pantothenate caused by overexpression of the *panF* gene product. This suggests that the transport of pantothenate does not regulate the biosynthesis of CoA.

In the yeast *S. cerevisiae*, the *FEN2* gene product is responsible for the transport of pantothenate. This gene product is a proton-coupled symporter with no similarity to the *panF* gene product of *E. coli* [87, 88]. Similarly, in *Schizosaccharomyces pombe*, the active protein in pantothenate transport is the *liz1* gene product, which is homologous to the *FEN2* gene product [89]. Both of these gene products are necessary for pantothenate uptake in low extracellular concentrations of the vitamin.

Many Gram-positive bacteria and Thermotogae utilise an energy-coupling factor (ECF) transporter for importing pantothenate [90]. This consists of a substrate binding S-component for pantothenate uptake (PanT), which associates with an ECF module to form a complete ECF transporter. Rodionov *et al.* [91] found the *panT* gene product mostly in members of the Firmicutes that are dependent on pantothenate for growth, such as *Streptococcus*, *Lactobacillales*, pathogenic members of the *Clostridia*, and *Mycoplasma*. Subsequently, the ECF-PanT of *Leuconostoc mesenteroides* has been shown to have specificity for transporting pantothenate after expression in a pantothenate transport-deficient, β-alanine-auxotrophic *E. coli* strain [92]. In addition to the pantothenate specificity, it was also shown that the activity of the transporter relies on the ECF module.

E. coli can reportedly make use of exogenous PantSH (**2.12**, Scheme 2.2) to fulfil its CoA biosynthetic requirements [93]. Additionally, 4'-phosphopantetheine (PPantSH, **2.10**), the phosphorylated form of PantSH as well as a product of both the turnover of the acyl carrier protein (ACP) prosthetic group and the degradation of CoA [94, 95], can be exported but not reimported by *E. coli*. Even though no specific transporters for PantSH and PPantSH have been identified, PantSH most probably travels into the cell via diffusion-based mechanisms due to its amphipathic character. This is also supported by numerous studies that have demonstrated the free uptake of a vast array of pantothenamide analogues in *E. coli* [96–98]. It has also been established that the growth of several bacteria can be supported by pantethine, the disulphide of PantSH, in media lacking pantothenate [99].



Scheme 2.2 The universal five-step CoA biosynthetic pathway (Black box) and the salvage pathway (Red box), with abbreviated enzyme names in bold and the corresponding enzyme-encoding genes in italics. Coenzyme A (**2.1**); Pantothenate (**2.2**); 4'-Phosphopantothenate (**2.8**); Phosphopantothenoyl-cysteine (**2.9**); 4'-phosphopantetheine (**2.10**); Dephospho-coenzyme A (**2.11**); Pantetheine (**2.12**). Adapted from Strauss [7].

2.2.2.3 Coenzyme A biosynthesis pathway

Following the biosynthesis or uptake of pantothenate (**2.2**), it is employed as the precursor for the universal CoA biosynthesis pathway (Black box, Scheme 2.2). This five-step pathway is the commonly used pathway in most organisms. In the first step of this pathway, 4'-phosphopantothenate (**2.8**) is produced by the phosphorylation of pantothenate, which is catalysed by pantothenate kinase (PanK). Phosphopantothenoylcysteine synthetase (PPCS) then catalyses the condensation of 4'-phosphopantothenate to form phosphopantothenoyl-cysteine (**2.9**). This is followed by phosphopantothenoylcysteine decarboxylase (PPCDC) that catalyses a decarboxylation of the cysteine moiety to yield PPantSH (**2.10**). The fourth step is catalysed by phosphopantetheine adenylyltransferase (PPAT), which involves the transfer of an ATP-donated adenylyl group to PPantSH that produces dephospho-coenzyme A (DePCoA, **2.11**). In the final step, DePCoA is phosphorylated by dephospho-coenzyme A kinase (DPCK) to produce CoA (**2.1**).

Alternatively, the salvage pathway (Red box, Scheme 2.2) can be utilised to synthesise CoA from PantSH (**2.12**) as a precursor. This pathway bypasses the second and third steps of the universal five-step pathway, which are catalysed by PPCS and PPCDC, respectively. Consequently, PantSH is phosphorylated by PanK to produce PPantSH, which can further be converted to CoA by the PPAT and subsequent DPCK-catalysed reactions.

2.2.3 Enzymes involved in the coenzyme A biosynthesis pathway

Even though there are many historical studies of the enzymes associated with the biosynthesis of CoA, the cloning and characterisation of all the enzymes in the pathway are reasonably recent events. In 1979, the gene encoding PanK in *Salmonella typhimuriun* was the first encoding gene of these pathway enzymes to be identified [100]. Subsequently, in 1992, PanK from *E. coli* was the first enzyme of CoA biosynthesis to be cloned and expressed [101]. The last enzyme of the pathway to be cloned was PPCS from *E. coli*, which occurred in 2001 [102]. However, thenceforth there has been immense development in the comprehension of the enzymes in this biosynthetic pathway.

2.2.3.1 Pantothenate kinase

PanK, the gene product of *coaA* or *coaX*, is the best studied of all the enzymes involved in the biosynthesis of CoA. After the first PanK (that of *E. coli*) was expressed and purified, the PanKs of several other organisms that represented mammal, plant, fungal, and eubacterial sources, followed [66]. It has recently been identified that PanKs, regardless of their catalytic function, can be subdivided into three main types, namely PanK type I, type II and type III. This differentiation is based on structural fold, sequence identity, as well as inhibition and

catalytic characteristics [103]. It was first assumed that the differentiation between the PanKs was along the phylogenetic lines, which noticeably defined prokaryotic and eukaryotic types. This notion of simple classification has since been discarded, as it has become evident that certain bacterial PanKs are closer related to previously assumed eukaryotic enzymes than to other bacteria [103]. In addition, some organisms possess two different PanK types, such as *Mycobacterium tuberculosis* [104].

PanK type I - The type I PanK of *E. coli* is the best representative of these enzymes as it was the first to be fully characterised and was regarded as the 'prototypical' bacterial PanK [101]. This enzyme is encoded by the *coaA* gene and exists as a homodimer [105, 106]. Furthermore, its structure suggests that the type I PanKs belong to the P-loop kinase family, which is a defining characteristic of this type of PanK [107, 108]. In addition to the highly cooperative binding of ATP, the binding of substrates also follow an ordered sequential mechanism, where ATP binds prior to pantothenate [106].

Another significant characteristic of type I PanKs is their particularly flexible substrate-binding sites. The ATP-binding site has been shown to allow competitive inhibition with the binding of CoA (and CoA thioesters, albeit less potent) [109, 110]. This facilitates feedback inhibition with the final product of the pathway, which is important for regulating intracellular CoA concentrations. Similarly, the low-level substrate specificity exhibited by the pantothenate-binding site enables PanK type I to phosphorylate various types of pantothenate analogues including PantSH [111].

PanK type II - The type II PanKs are mostly harboured by eukaryotic organisms, such as plants [112, 113], fungi [114], mammals [115, 116], and insects [117]. Several of these organisms possess numerous PanK-encoding genes, where each gene is expressed in a specific tissue. In humans, for example, there are four main encoding genes, namely *PANK1*, *PANK2*, *PANK3*, and *PANK4* [66]. *PANK1* is mainly expressed in the kidney and liver and less in the muscle and heart, while *PANK2* is expressed in most tissues and localises in the mitochondria [118, 119]. Conversely, *PANK3* is highly expressed in the liver and to a lesser extent in the skeletal muscle and heart, whereas *PANK4* expression is located in the heart and muscle [120]. Moreover, some of these genes can encode two related but different PanK isoforms. Examples of this include *PANK1* of humans and mice, which has PanK1*α* and PanK1*β* as protein products [121, 122].

Type II PanKs do share some similarities with type I PanKs with regards to substrate specificity and inhibition. The type II PanKs can also undergo feedback inhibition by CoA biosynthesis end products. However, in type II PanKs, the CoA thioesters (acetyl-CoA in

particular) are usually more potent inhibitors, compared to free CoA [116, 122–124]. Similar to type I PanKs, these PanK types also accept pantothenamides (amide containing analogues of pantothenate) as substrate alternatives [125]. Nevertheless, type I and type II PanKs differ in their primary sequence profiles and structural folds. The type II PanKs are not members of the P-loop kinase family like their type I counterparts. In fact, the structures of various type II PanKs suggest that these enzymes belong to the sugar kinase/heat-shock protein 70/actin (ASKHA) superfamily of kinases [107, 126, 127].

Additionally, it has been demonstrated that some Gram-positive bacteria possess PanK-encoding genes that are related more closely to the 'eukaryotic' type II PanKs (described above) than to the 'prototypical' bacterial type I PanKs, based on primary sequence homology. The enzyme of *Staphylococcus aureus* is one of these bacterial type II PanKs [128, 129]. This enzyme, however, is not affected by feedback inhibition by CoA and its thioesters like its eukaryotic counterparts [128, 129]. Yet, in spite of this, the PanK of *S. aureus* still accepts pantothenamides as substrate alternatives [125, 128].

PanK type III - The suggestion that a third PanK type exists first arose following genome-wide analyses where gene homologues encoding all the CoA biosynthetic enzymes, except PanK, were identified in some bacteria, such as *Pseudomonas aeruginosa* and *Helicobacter pylori* [83, 130]. The results remained the same irrespective of submitting type I PanK or any type II PanK in the homology searches. Following this, a PanK-encoding gene, distinguishable from other PanK types, was reported in *Bacillus subtillus* [131]. This gene was termed *coaX* in order to differentiate between it and the more popular *coaA* gene. Subsequently, homology searches based on the *coaX* sequence identified homologues in other bacteria that were initially thought to have apparent missing PanK enzymes [66]. Interestingly, this gene is exclusively present in a broad range of pathogenic bacteria [132].

After the expression, purification and full characterisation of the CoaX protein of *B. subtillus* and *H. pylori*, it was revealed that these PanKs demonstrated a noticeable difference in kinetic parameters compared to that of the other PanK types, in addition to the variation in primary sequence and structure [133]. For instance, even though the affinity for pantothenate of these enzymes were relatively the same as that of the other PanKs, their K_m values for ATP were reported to be in the millimolar range (up to ~10 mM for *H. pylori*) compared to the micromolar ranges for type I & II PanKs. Furthermore, it was also shown that these PanKs do not experience feedback inhibition by CoA or acetyl-CoA nor do they accept pantothenamides as alternative substrates [103, 127, 133]. Thus, PantSH can also not be accepted as a substrate, which implies that the salvage pathway (Red box, Scheme 2.2) is not an option in the organisms containing a type III PanK. On the basis of

these attributes, the CoaX proteins were recognised as the third type of PanK, and thus named type III PanKs. Moreover, structural studies have established that the type III PanKs are also members of the ASKHA superfamily like their type II counterparts [103, 127, 132].

2.2.3.2 Phosphopantothenoylcysteine synthetase

Two main forms of PPCS have been identified to date, specifically bacterial PPCS and eukaryotic PPCS. The bacterial PPCS utilises CTP to activate the carboxylate of 4'-phosphopantothenate (**2.8**, Scheme 2.2) by forming an acyl-cytidylate intermediate [102]. This bacterial version of PPCS typically forms a fusion protein with the subsequent enzyme in the pathway, PPCDC, which produces the bifunctional CoaBC protein (protein product of *coaBC*). Conversely, the eukaryotic form of the enzyme (protein product of *coaB*) is monofunctional and utilises ATP to activate the substrate carboxylate [134–136]. This also seems to be the case in *Streptococci* and *Enterococci*, as they are predicted to be exceptions to the bacterial bifunctional enzyme formation [130]. However, both the eukaryotic and bacterial PPCS exhibit dimeric structures with similar folds [137].

Following the expression, purification and characterisation of the bacterial CoaBC protein, it was reported to have a K_m value of 55 µM for 4'-phosphopantothenate, with a K_m value of 106 µM and 109 µM for CTP and L-cysteine, respectively [102, 138]. Studies have also shown that the product of PPCS in the CoaBC fusion protein, phosphopantothenoylcysteine (**2.9**, Scheme 2.2), dissociates from the protein and then binds to a separate active site on the PPCDC domain to commence further transformation [139]. This suggests that the merging of these two enzymes in bacteria offer no extensive functional significance. This fusion does, however, ensure that the concentration of the PPCS and PPCDC enzymes remain equal [139].

PPCS has also been shown to have a very high selectivity for cysteine [140]. In a study by Strauss *et al.* [140], the enzyme was exposed to 500 000-fold excess serine, which has a structure that closely resembles that of cysteine, but only cysteine was integrated into the product. This selectivity for cysteine is not surprising, as the structure of serine has a thiol group instead of the hydroxyl group from cysteine. Therefore, if the condensation of serine proceeds, it could result in the formation of oxy-CoA, a potential toxic CoA analogue [141].

2.2.3.3 Phosphopantothenoylcysteine decarboxylase

As mentioned previously, prokaryotic PPCS and PPCDC enzymes fuse to form a bifunctional CoaBC protein product, with the exceptions of *Streptococci* and *Enterococci*, which supposedly possess monofunctional enzymes. These prokaryotic CoaBC enzymes share a structural feature with the lantibiotic decarboxylases, which is that they form

homododecamers (tetramers of trimers) [142, 143]. This occurs via interactions of the CoaC domains. However, the eukaryotic PPCDC enzymes are reported to be monofunctional proteins with a trimeric structure [144].

The activity of the PPCDC enzyme has proved to be far more elusive than all the other enzymes in the CoA pathway, concerning both its mechanism and identification. The enzymatic mechanism was finally solved in 2000, after the characterisation of the PPCDC domain associated with the *E. coli* CoaBC protein [143]. The reaction catalysed by PPCDC causes the formation of a negative charge on the carbon next to the amide of phosphopantothenoylcysteine. This charge is then stabilised by a flavin mononucleotide cofactor (tightly bound to the enzyme), which is responsible for the oxidation of the substrate cysteine thiol. Subsequently, a thioaldehyde is formed, followed by its spontaneous decarboxylation to produce an enethiol, which is then reduced back to yield PPantSH (**2.10**, Scheme 2.2). This last step relies on an active site cysteine, which was reported to be absent in the CoaBC protein of *Methanocaldococcus jannaschii* and instead, it was replaced by a glutamic acid [145]. Consequently, this indicated that PPCDC enzymes may exhibit variations in their mechanisms.

2.2.3.4 Phosphopantetheine adenylyltransferase

PPAT, also known as the CoaD protein (protein product of *coaD*), is the enzyme responsible for catalysing the penultimate step in the CoA biosynthetic pathway. In bacteria, PPAT is reported to be a single enzyme [146]. This enzyme adopts a homohexameric structure, which consists of a dimer of separate trimers [147–150]. Following a comprehensive kinetic characterisation of the *E. coli* PPAT, it was revealed that this enzyme follows an arbitrary bi-bi mechanism in which a ternary complex of ATP, PPantSH and enzyme is formed [151]. The K_m values that were determined are 4.7 ± 0.5 µM for PPantSH and 220 ± 10 µM for ATP. In addition, it was also reported that the bacterial PPAT experiences feedback inhibition. DePCoA (**2.11**, Scheme 2.2) binding was revealed to be competitive with both PPantSH and ATP binding; whereas the binding of pyrophosphate was competitive with ATP (pyrophosphate product inhibition of PPantSH binding could not be determined). The binding of CoA is also competitive with PPantSH, ATP and DePCoA. Furthermore, CoA thioesters are much weaker inhibitors compared to free CoA, suggesting that CoA is the main inhibitor in feedback inhibition of this enzyme. According to structural studies of bacterial PPAT, they exhibit sequence homology to members of the nucleotidyltransferase *α*/*β* phosphodiesterase enzyme superfamily [148].

In contrast, the PPAT activity in most eukaryotes is fused to DPCK, which is the last enzyme in the pathway [152]. This produces a bifunctional PPAT/DPCK protein, better known as

CoA synthase (CoASy). The sequences of these eukaryotic bifunctional proteins have been shown to have little similarity with that of the monofunctional bacterial enzymes [134, 135, 153, 154]. There is also little known about the structure of the PPAT domain in the CoASy protein. However, preliminary studies propose that they too belong to the superfamily of nucleotidyltransferase *α*/*β* phosphodiesterases [155].

2.2.3.5 Dephospho-coenzyme A kinase

As mentioned previously, the DPCK activity in eukaryotes is fused to PPAT as part of the CoASy bifunctional protein [152]. However, in contrast to PPAT, the sequence of the DPCK domain of CoASy demonstrates good homology with its prokaryotic counterpart, which is reported to be a monofunctional protein [135, 153, 154]. The expression and characterisation of the *E. coli* DPCK protein (protein product of *coaE*) revealed that it is a 22.6 kDa monomer in solution with K_m values reported being 740 µM and 140 µM for DePCoA and ATP, respectively [156]. Conversely, the DPCK domain of the CoASy protein demonstrated a similar affinity for ATP (reported K_m value of 192 µM) but a much higher affinity for DePCoA (reported K_m value of 5.2 ± 1.5 µM) [154]. This results in the irreversible PPAT activity of the bifunctional protein, which is normally reversible as a monofunctional protein. The structural analyses of numerous DPCK enzymes, including the monofunctional protein of *E. coli* and the CoASy DPCK domain of *Mus musculus*, revealed that they belong to the P-loop kinase enzyme family [107, 157].

There are still some uncertainties regarding the kinetic parameters of DPCK, due to the limited information available on this enzyme. Although very little is known about the catalytic mechanism of DPCK, it has been proposed that a conformational change may take place during catalyses. The crystal structures of *E. coli* DPCK (bound to ADP only) and *Thermotoga maritima* DPCK (bound to DePCoA and ADP) were superimposed to reveal a distinct difference in conformation between the two enzymes [157]. This change suggested that the enzyme adopts a closed conformation after binding to DePCoA in order to bring it closer to ATP catalytic activity. In fact, similar substrate-induced conformational changes are not uncommon in the P-loop kinase enzyme family [158].

If this conformational change-hypothesis is true for *E. coli* DPCK, it may explain the low substrate affinity for DePCoA observed in the *in vitro* studies, as the initial binding of the substrate has a low affinity, followed by the conformational change to a structure with a higher affinity to allow catalysis to proceed. Alternatively, since the DPCK of *E. coli* has been shown to be a monomer [156] and yet in the presence of sulphate it crystallises as a trimer [157]; the apparent low affinity observed for DePCoA could be due to the changes in quaternary protein structure.

2.2.4 Coenzyme A biosynthesis pathway as a drug target

CoA biosynthesis has long been considered as a potential target for the development of antimicrobial agents [8]. There are several reasons for this: first, the importance of CoA and its thioesters for the survival of all living organisms is undeniable; second, most organisms are believed to obtain CoA via *de novo* biosynthesis [130]; and third, even though the biochemistry of CoA biosynthesis is relatively conserved, there are significant differences in the sequence and structure of the biosynthetic enzymes in prokaryotes compared to their eukaryotic counterparts [66], which suggests that selective inhibition of microbial CoA biosynthesis should be possible. Therefore, most of the enzymes in the CoA biosynthetic pathway are excellent potential targets for developing selective inhibitors. PPCS and PPCDC, however, are possible exceptions, since these enzymes may be bypassed by organisms with the ability to utilise the salvage pathway.

2.2.4.1 Pantothenate kinase inhibition

Of all the enzymes associated with the biosynthesis of CoA, PanK has received the most attention as a target for developing antimicrobial agents [6]. This can be attributed to three main reasons: (a) PanK catalyses the first committed step in the CoA biosynthetic pathway; (b) it is also believed to catalyse the rate-limiting step in this pathway; and (c) PanK enzymes show remarkable diversity by exhibiting three distinct identifiable types, of which type II occurs predominately in eukaryotes. These reasons suggest that selective inhibition of PanK, and ultimately CoA biosynthesis, in pathogenic micro-organisms should be possible without affecting the enzyme of their host.

The majority of the inhibition studies done on PanK type I have involved the pantothenamide class of pantothenate analogues. However, these pantothenamides are strictly speaking not inhibitors of PanK as they do not inhibit the activity of the enzyme. Instead, they act as alternative substrates producing anti-CoA metabolites [96]. These pantothenamide compounds have an *N*-substituted amide as opposed to the carboxylic acid of pantothenate. The compounds with *N*-heptyl and *N*-pentyl substituents are some of the best-characterised examples of such alternative substrates [96, 129, 159], even though various other substituents have also been prepared [125]. The type I PanKs frequently display k_{cat}/K_m values for pantothenamides that are similar or even higher than for pantothenate or PantSH [96]. Thus, these PanK enzymes do not discriminate between them catalytically, which leads to a decrease in the phosphorylation rate of pantothenate in the presence of these compounds. Nevertheless, these pantothenamide compounds only cause inhibition once they are accepted as alternative substrates. However, this ability to accept alternative substrates is only a characteristic of some types of PanKs. Consequently, if for example, an

organism possesses a PanK with higher substrate specificity, the use of pantothenate analogues for inhibition or production of anti-CoA metabolites will be ineffective. Several direct inhibitors of *Mycobacterium tuberculosis* PanK have, however, been identified but none have shown whole cell activity against the organism when tested under *in vitro* conditions [160–162].

The work done on PanK type III inhibitors have been far less since it was discovered quite recently [133]. The only known inhibitors of this type of PanK are nucleoside triphosphate mimetics of ATP. These inhibitors were discovered by preparing a library of ATP structural analogues, in which uncharged methylene-triazole-linked monosaccharide groups replaced the triphosphate side chain of ATP, for the inhibition of *Bacillus anthracis* PanK [163]. Of these ATP analogues, one exhibited inhibition of the *B. anthracis* PanK by competing with ATP binding, with a reported K_i value of 164 µM, which is significantly lower than the K_m value for ATP (510 µM). Even though this suggests that the enzyme's affinity for the inhibitor is much higher than for ATP, the K_i value is still not low enough to be of pharmaceutical interest. Furthermore, no whole cell inhibition was reported. This low affinity for ATP demonstrated by PanK type III has also raised some questions on whether or not ATP is the co-substrate for these enzymes [127, 133]; and in turn, is the ATP mimetics then the best approach for going forward in search of PanK type III inhibitors. Nonetheless, the search for PanK inhibitors will continue, as no inhibitors for PanK activity have been identified to exhibit inhibition of cell growth to date [6].

2.2.4.2 Phosphopantetheine adenylyltransferase inhibition

Due to the reported regulatory role of PPAT in CoA biosynthesis [151] and the little sequence similarity the eukaryotic PPAT domain of CoASy shares with the bacterial PPAT enzyme [134, 153, 154], it is another enzyme in the pathway that has received special attention as a potential target for developing selective antimicrobial agents. Thus, numerous inhibitors have been identified, although none of these have demonstrated any inhibition of whole cell growth [164, 165]. However, in a recent study, PPAT was validated as a target for antimicrobial therapy by implementing a high-throughput screening of an AstraZeneca compound library [166]. This study described a variety of cycloalkyl pyrimidines as possible inhibitors of both Gram-negative and Gram-positive bacteria. Subsequently, structure-based optimisation of these compounds resulted in potent inhibitors against Gram-positive bacteria, which exhibited competitive inhibition of *Streptococcus pneumoniae* and *S. aureus* PPAT with regards to PPantSH. Despite the fact that these inhibitors demonstrated effective growth inhibition of several clinical Gram-positive isolates, their biological activity could not

be reconciled with drug-like properties and therefore cannot be considered as suitable clinical candidates.

In another recent study, lead compounds were identified as targets for *H. pylori* PPAT by applying an *in silico* screening of the crystal structure of this enzyme with 407 compound structures obtained from the PubChem compound database [167]. Consequently, D-amethopterin (methotrexate), a known anti-cancer drug and antimetabolite, was identified as a potential inhibitor of *H. pylori* PPAT. This compound demonstrated inhibition of *H. pylori* growth, as well as inhibition of the PPAT enzyme activity *in vivo* by preventing the binding of both ATP and PPantSH. Unfortunately, low potency was reported in both cases. Therefore, the authors suggested that further structure optimisations to methotrexate may increase interactions with *H. pylori* PPAT. In addition, the target selectivity would also have to be investigated.

2.2.4.3 Dephospho-coenzyme A kinase inhibition

As of yet, DPCK has been overshadowed by its fellow CoA biosynthesis enzymes regarding its potential as a drug target and has therefore not been the focus of many development studies for selective inhibitors [6]. This is mostly due to the high sequence and structure homology shown by the bacterial DPCK enzyme and the DPCK domain of the human CoASy enzyme, which suggests that selective inhibitors of the bacterial enzyme will be difficult to develop [135, 153, 154]. However, the reportedly low substrate affinity of *E. coli* DPCK might be an indication that this enzyme can accept alternative substrates, which can lead to either inhibition of the catalytic activity or production of anti-CoA metabolites.

A recent review of the results obtained from AstraZeneca's antibacterial discovery efforts between 2001 and 2010 reported that one of the highlighted high-throughput screenings was for *S. pneumoniae* DPCK antibacterial agents [168]. This screening led to the identification of compound structures belonging to the cofactor mimetic antibacterial class. Unfortunately, there are no published cellular studies in which the predicted compound structures are evaluated for inhibition of the anticipated target enzyme within intact cells.

2.2.5 Coenzyme A biosynthesis in mycoplasmas

Thus far, there have not been many studies in which the biosynthesis of CoA in mycoplasmas is extensively investigated. This lack of information on this pathway might be due to the small genome and subsequent absence of various biosynthetic pathways in mycoplasmas, which allowed assumptions to be made that the pathway is not essential for survival and therefore was not investigated. Another reason for this might be attributed to the

poor annotation (or in some cases insufficient sequencing) of most mycoplasma genomes, which can lead to the presumption that the genes are absent.

Nonetheless, most of the existing studies report that mycoplasmas lack all the enzymes in the biosynthetic pathway apart from DPCK [8, 135]. However, there are other studies that disagree with this, stating that some mycoplasmas indeed have a PPAT and a DPCK, namely *M. synoviae*, *M. mycoides*, *M. mobile*, *M. pulmonis* and *M. penetrans* [169], but still, no studies have reported the presence of the other pathway enzymes. This inconsistency emphasises the need for a comprehensive investigation into the CoA biosynthesis of mycoplasmas in order to determine its suitability as a potential target for drug or vaccine development in mycoplasmas.

Chapter 3 – Bioinformatics analysis of coenzyme A biosynthetic pathway enzymes in mycoplasmas

3.1 Introduction

As the demand for ostrich products (such as leather, feathers, eggs, and meat) increases, more intensive rearing strategies have to be applied. Consequently, this generates a vulnerable environment, which is beneficial for the spreading of pathogens and subsequently diseases. Ostrich-infecting mycoplasmas, associated with respiratory infections, are a serious concern in this environment as they can lead to major production losses. Currently, three *Mycoplasma* species associated with ostrich infections have been identified: Ms01, Ms02 and Ms03 [5]. Disease management is highly important for the conservation of the ostrich industry, which is also a key contributor to the South African economy. Thus, treatment against these mycoplasma infections is vital.

In an attempt to identify novel targets for potential treatment development, the coenzyme A (CoA) biosynthetic pathway has been selected. The universal CoA biosynthetic pathway consists of five enzyme-assisted steps (Figure 1.1); of which the second and third steps can be bypassed in some organisms to employ an alternative three-step salvage pathway [7]. This biosynthetic pathway is well-known for its potential as a drug target [8], yet in mycoplasmas, it is relatively unexplored. Hence, before the CoA biosynthetic pathway can be considered as a suitable target in mycoplasmas, it must first be investigated. The aim of this study was therefore to investigate the enzyme-encoding genes associated with the CoA biosynthesis pathway in *Mycoplasma* species by using a bioinformatics approach. Firstly, currently annotated mycoplasma genomes were examined to determine the presence or absence of annotated genes that encode for each of the five enzymes involved in CoA biosynthesis. In genomes where one or more of these genes were not annotated, their presence or absence was determined using BLAST (Basic local alignment search tool) analysis. To confirm the identity of the newly found gene homologues, the functional motifs and domains of their protein products were determined. Finally, the phylogenetic relationships of these *Mycoplasma* species were determined using 16S rRNA sequence data and protein sequence data of the respective CoA biosynthetic enzymes to determine the functional relationships of these proteins between species.

3.2 Background on phylogenetic approaches

Phylogenetic analyses are used to address various questions regarding epidemiological dynamics of pathogens, evolutionary histories of populations, and relationships between

genes or species [170]. The relationships between species can be determined by making use of a marker gene. This gene must be present and have the same function in all of the species in the assessment besides being under the same selective pressure. Therefore, the selection of the marker gene is crucial.

The 16S rRNA gene was the chosen marker gene in this study. This gene is highly conserved and universally distributed amongst most species of archaea and bacteria [171]. The structural role of the 16S rRNA genes in ribosomes, which function in all self-replicating cells during protein synthesis, is the same regardless of species [172, 173]. Furthermore, the 16S rRNA genes are under similar evolutionary pressure. Even though other marker genes have been used, the 16S rRNA gene is currently recommended for phylogenies of species in the class Mollicutes [25, 26]. Based on this gene, the class Mollicutes can be divided into five distinct phylogenetic groups, namely the hominis-, pneumoniae-, spiroplasma-, anaeroplasma-, and asteroleplasma groups (Figure 2.1) [25, 29, 30]. The *Mycoplasma* species are distributed amongst the former three.

A phylogeny of the individual enzymes associated with the biosynthesis of CoA was also constructed in this study to analyse the functional relationship amongst different copies of the respective enzymes between species. Although the genes that code for these enzymes do not necessarily fall under the same selective pressure, they are subjected to restrictions of protein structure and function. The amino acid sequences were used for these phylogenies, as the nucleotide sequences are very heterogeneous and could, therefore, not be aligned with confidence.

The online alignment program Clustal Omega [174, 175], was used for the sequence alignments of both the 16S rRNA gene sequences and the respective enzyme amino acid sequences. Clustal Omega is a multiple sequence alignment (MSA) program that can be used for the alignment of divergent sequences. It uses an iterative progressive alignment approach with hidden Markov models in combination with HHalign [176] for increased accuracy. Clustal Omega also uses a modified version of the mBed algorithm [177] to generate its guide trees, which is a much faster method than the traditional approach of calculating full distance matrices. In addition to its fast runtime, Clustal Omega also exhibits similar accuracy levels when compared to the consistency heuristic programs, such as T-Coffee and Probcons [174].

The construction of phylogenies is either character-based or distance-based [178]. In the latter methods, namely neighbour-joining, distances are calculated between pairs of sequences within the complete alignment [179]. This generates a distance matrix, which is

then used to construct a phylogenetic tree. Although these distance-based analyses are computationally efficient, they are very sensitive to gaps in the alignment and perform quite poorly with divergent sequences [178]. Conversely, character-based methods simultaneously examine all sequences in the alignment one character (a site or column in the alignment) at a time to calculate a score for each possible tree. The tree with the best score is then retained. Parsimony, Bayesian inference and maximum likelihood are all examples of this type of method. The “tree scores” in parsimony are the minimum number of changes; in Bayesian inference, the posterior probability; and in maximum likelihood, the log-likelihood values [180, 181]. Additionally, each of these respective analyses uses distinctive algorithms, formulas and models that contribute to their outcome.

Currently, it is widely accepted that statistical methods, such as maximum likelihood [182], generate more reliable results compared to distance-based and parsimony methods [178, 183]. One major shortcoming of a parsimony analysis is its lack of explicit assumptions, which makes it very difficult to integrate knowledge of sequence evolution into tree reconstruction. This is not the case with maximum likelihood, as all the model assumptions are explicit and can, therefore, be evaluated and improved upon. The main drawback of maximum likelihood methods is that they tend to be computationally more demanding [184]. However, this can be overcome by utilising online platforms, such as CIPRES Science Gateway [185]. One of the programs available on the CIPRES Science Gateway web portal is RAxML (Randomised Axelerated Maximum Likelihood) [186], a maximum likelihood program that can be used for phylogenetic analysis of large datasets. This program implements the standard Subtree-Pruning-and-Regrafting (SPR)-based hill-climbing algorithm and applies various important heuristics to reduce the amount of unfavourable SPR candidates.

3.3 Materials and Methods

3.3.1 Selection of *Mycoplasma* species used in the investigation

A total of 62 *Mycoplasma* species were investigated in this study (Supplementary Table 1.1) of which some species have a complete genome available and others have genomes available only on contiguous sequence (contig) level, however all genomes were annotated. Using the 16S rRNA phylogeny of Wium *et al.* [45], a group of 28 *Mycoplasma* species, including the ostrich-infecting *Mycoplasma* sp. Ms02 (Ms02), were selected that represented different 16S rRNA phylogenetic groupings. The selected species were those that had CoA biosynthesis pathway information available on the KEGG pathway and/or SEED viewer subsystems databases, apart from Ms02 whose genome was annotated using the RAST

prokaryotic genome annotation server, which is currently not publically available [57]. Ten additional species, not represented in the phylogeny of Wium *et al.* [45], were also included since their genomes were available on the KEGG pathway and/or SEED viewer subsystems databases. The remaining species were added following BLAST searches, which were restricted to the class Mollicutes, in order to obtain a larger species variety for the investigation (as discussed in the following sections). It was decided to include only one strain of a selected species to avoid redundancy.

3.3.2 Identification of currently annotated CoA biosynthesis enzyme-encoding genes in mycoplasma genomes according to the KEGG pathway, SEED viewer subsystems and NCBI databases

Both the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway database and the SEED viewer subsystems database were used as a source of metabolic pathway information. KEGG is a database resource that integrates genomic, systemic, chemical, and health information provided by 15 manually curated databases and one computationally generated database (16 in total) [187–189]. Thus, gene catalogues in fully sequenced genomes are integrated with higher-level systemic functions of the cell and organism, such as metabolism and other cellular processes. The SEED viewer subsystems database, however, was developed as a platform for annotating genomes based on subsystems [190]. Therefore, this database consists of populated subsystems developed at other sites, along with those obtained from annotated and curated genome submissions.

The CoA biosynthetic pathway was selected as the pathway of interest in both the KEGG pathway and SEED viewer subsystems databases, followed by the identification of annotated enzyme-encoding genes of this pathway within the selected genomes of *Mycoplasma* species available to the respective database. The results of both databases were manually compared and subsequently combined. The amino acid sequences of the identified enzymes were obtained from the National Centre for Biotechnology Information (NCBI) database [191, 192]. In addition, the nucleotide and corresponding amino acid sequence of the Ms02 dephospho-CoA kinase (DPCK) enzyme was obtained from its previously annotated genome [57].

BLAST is a sequence similarity search platform that is used to compare a user-specified query sequence to a database of sequences (in this case the NCBI database) [193, 194]. In order to identify currently annotated enzyme-encoding genes of the CoA biosynthetic pathway in genomes of *Mycoplasma* species that were not available on the KEGG and SEED viewer databases, a protein-protein BLAST (BLASTP) was conducted using the

amino acid sequences obtained from annotated genomes (described in the previous section) as query. The BLASTP search was restricted to the class Mollicutes (taxid: 31969).

Furthermore, Position-Specific Iterated BLAST (PSI-BLAST) searches [195, 196] were performed using the same amino acid sequences as in the BLASTP search as query to identify additional currently annotated enzyme-encoding genes that could not be identified with BLASTP. PSI-BLAST generates a profile (basically a consensus sequence) of significant user-specified BLAST matches, which is then used as query in the following BLAST iteration. For both the BLASTP and PSI-BLAST searches, the sequence of a hit was only included for further analysis if the sequence of the hit could be associated with a mycoplasma of which the genome information was available.

3.3.3 Identification of CoA biosynthesis enzyme-encoding gene homologues

In the *Mycoplasma* genomes previously selected, there were some in which one or more of the five CoA biosynthetic enzyme-encoding genes were not annotated. In such instances, the same method as described in the previous section was used in order to identify enzyme-encoding gene homologues. If enzyme-encoding gene homologues could not be identified in a specific genome using these strategies, the BLASTP search was repeated using the amino acid sequences from annotated genomes as query (as before), but this time the BLAST search was restricted to the specific *Mycoplasma* species in which a gene homologue was to be identified. Additionally, gene homologues in the genomes of *Mycoplasma* species, which were not amongst the previously selected species, were also identified using the above-mentioned strategies.

During the annotation of sequences, genes with uncertain properties can be categorised according to reliability and are annotated as probable, potential, putative or hypothetical in order of decreasing reliability [197–199]. However, in this study an identified sequence in the database was deemed a homologue if it exhibited an amino acid identity of at least 30% over at least 70% of the query sequence length.

3.3.4 Analysis of genome locations of enzyme-encoding genes and the association between number of genes and energy source used by species or its host

The amino acid sequences of the BLAST-identified genes were obtained from the NCBI database and the genomic locations of these genes, in their respective genomes, were determined. This was done to possibly assist with confirming the identity of newly found homologues. Furthermore, an analysis was done to determine if there is a correlation

between the number of enzyme-encoding genes within the CoA biosynthetic pathway of a mycoplasma and its preferred substrate source for general energy metabolism or the host of the species.

3.3.5 Bioinformatics analysis of annotated and newly identified CoA biosynthesis enzyme-encoding gene homologues

Bioinformatics analyses on both the annotated and newly identified amino acid sequences of all the pathway enzymes were performed to confirm the identity of homologues as CoA biosynthetic enzymes. Initially, NCBI's Conserved Domain Database (CDD) v3.16 was used to identify conserved domains and family relationships in the protein sequences [200, 201]. Hits with an E-value >1.0e-05 were disregarded. InterPro v65.0 was also used for additional verification, as it uses signatures supplied by a combination of 14 different member databases, including PROSITE, ProDom, SUPERFAMILY and Pfam [202]. Hence, by using InterPro, these databases can be scanned simultaneously.

Furthermore, conserved motifs within the protein sequences of each enzyme were determined using the Multiple Expectation maximisation for Motif Elicitation (MEME) algorithm from the MEME suite v4.12.0 [203, 204]. This algorithm does not rely on existing database information for motif discovery; it simply identifies sequence similarities amongst the submitted group of sequences. Thus, the identified motifs may not necessarily correspond to a specific function. The search parameters were set to identify motifs of a length between 6 and 50 amino acids with a maximum of 4 motifs per sequence. The predicted motifs were compared to known functional motifs of the corresponding enzymes.

3.3.6 Multiple sequence alignments and phylogenetic analyses

In order to determine the evolutionary relationship between the *Mycoplasma* species used in this study, a phylogenetic analysis based on the 16S rRNA gene was performed. Subsequently, the 16S rRNA nucleotide sequences of the 62 *Mycoplasma* species (Supplementary Table 1.1), as well as a representative species of closely related genera (namely *Clostridium, Lactobacillus, Streptococcus,* and *Bacillus*), were obtained from the NCBI database. The chosen representatives used were: *Clostridium innocuum* (KR364751.1), *Lactobacillus fermentum* (FJ462686.1), *Streptococcus pneumoniae* (NR_115239.1), and *Bacillus coahuilensis* (EF014447.1); of which the latter was used as outgroup. These four genera were chosen since mycoplasmas are believed to have originated from a node in a branch of low G+C containing gram-positive bacteria (based on 16S rRNA sequence data) in which species of these four genera are found. Thus, by including these representative species, the investigated mycoplasmas and their proposed

ancestors can be compared. *B. coahuilensis* was, subsequently, chosen as outgroup since it is the most distantly related species amoungst these four representatives.

Additionally, a phylogenetic analysis based on each of the enzymes associated with the CoA pathway was performed to investigate the functional relationship between the respective proteins. The previously obtained amino acid sequences (Supplementary Table 1.1) were used in this analysis, along with amino acid sequences of the respective enzymes of the same representative species of closely related genera that were used in the 16S rRNA gene phylogeny. The amino acid sequences of the *B. coahuilensis* proteins were used as outgroup in all of the phylogenetic analyses.

All of the multiple sequence alignments (nucleotide and amino acid sequences) were done using Clustal Omega v1.2.3 [174, 175], which was then further refined manually using BioEdit v7.0.5.2 [205]. If required the Geneious v11.0.2 [206] program was utilised because of its enhanced alignment functionalities, which allowed further optimisation of MSAs.

Maximum likelihood phylogenies were constructed using RAxML-HPC2 on XSEDE v8.2.10 [186] via the CIPRES Science Gateway v3.3 [185] web portal. Evaluation of clade support was performed with Bootstrap analysis, which was set to automatically stop when the majority rule criterion is reached, according to program recommendations. Bootstrap values ≥75% were considered resolved and well supported, values ≤75% but ≥50% were considered resolved but moderately supported, and values ≤50% were considered poorly supported. Thus, only values 50% and above are indicated on the phylogenetic trees.

3.4 Results

3.4.1 Identification of currently annotated CoA biosynthesis enzyme-encoding genes in mycoplasma genomes according to the KEGG pathway, SEED viewer subsystems and NCBI databases

According to the KEGG pathway and SEED viewer subsystems databases, 25 of the 37 *Mycoplasma* species (excluding Ms02) have at least one of the five CoA biosynthetic pathway enzyme-encoding genes (Table 3.1). Of the 25 *Mycoplasma* species, *Mycoplasma gallinaceum* was the only species that was identified to possess no DPCK gene. However, according to the KEGG database, *M. gallinaceum* does possess the other enzyme-encoding genes. *Mycoplasma mobile* and *Mycoplasma synoviae* are the only two of the 25 species that were identified to possess all five of the enzyme-encoding genes. Together with *M. gallinacium*, they were the only three species that, according to the KEGG pathway and SEED viewer subsystems databases, contained a gene for phosphopantothenoylcysteine

synthetase (PPCS) and a gene for phosphopantothenoylcysteine decarboxylase (PPCDC). However, following the download of the amino acid sequences from NCBI, it was revealed that these species (excluding *M. gallinacium*) possess the *coaBC* gene that codes for a bifunctional phosphopantothenoylcysteine decarboxylase/phosphopantothenoylcysteine synthetase (CoaBC) enzyme. The genes that were reported by the KEGG pathway database to code for PanK and CoaBC in *M. gallinacium* were annotated as hypothetical proteins in the NCBI database.

In addition, the PanK-encoding gene that was present in six of the 25 *Mycoplasma* species was identified to be the type III PanK enzyme. Four of these six species were identified to possess no PPCS or PPCDC, yet they possess a phosphopantetheine adenylyltransferase (PPAT) and DPCK (the three enzymes involved in the salvage pathway). This is unorthodox since the type III PanKs cannot accept pantetheine as a substrate, which is the precursor for the salvage pathway.

Furthermore, 12 *Mycoplasma* species available on the KEGG database were identified to possess none of the CoA biosynthetic pathway enzyme-encoding genes. There are also several species for which the two databases differ regarding the available information on the CoA biosynthetic pathway.

A summary of the CoA biosynthetic pathway enzymes of all 62 *Mycoplasma* species that were investigated is given in Table 3.2. Of the 62 species, 53 species had at least one annotated CoA biosynthetic enzyme-encoding gene. Amongst these, the DPCK-encoding gene was the only annotated gene that could be identified in all of the investigated *Mycoplasma* species, except for *M. gallinaceum*. Moreover, there were seven species identified to have an annotated bifunctional CoaBC enzyme-encoding gene and one species (*Mycoplasma iowae*) with separate annotated PPCS and PPCDC enzyme-encoding genes.

Moreover, all of the identified PanK genes encode for a type III PanK, of which six species were identified with the unorthodox combination of three annotated salvage pathway enzyme-encoding genes (PanK, PPAT and DPCK). Of the investigated *Mycoplasma* species, there were 23 species in which only an annotated PPAT- and DPCK-encoding gene could be identified and another six species in which only an annotated DPCK-encoding gene could be identified. The investigation also revealed five species that possess an unknown bifunctional haloacid dehalogenase (HAD)-like protein/dephospho-CoA kinase (HAD-DPCK) protein, which was annotated as DPCK. In four of these five species, this bifunctional HAD-DPCK protein was the only CoA biosynthetic pathway enzyme that could be identified.

However, in the fifth species (*Mycoplasma conjunctivae*), an annotated PPAT-encoding gene was also identified.

Additionally, there does not seem to be a correlation between the number of identified CoA biosynthetic enzyme-encoding genes that are present in the species and the 16S rRNA phylogenetic grouping of the species (Table 3.2).

Table 3.1 Annotated coenzyme A biosynthetic pathway enzymes identified in *Mycoplasma* species available on the KEGG pathway and SEED viewer subsystems databases

Phylogenetic group[a]	*Mycoplasma* species	PanK	CoaBC*	PPAT	DPCK
Homins group	*M. agalactiae*			1	1
	M. alligatoris	3		3	3
	M. bovis			1	1
	M. californicum			2	2
	M. canis			2	2
	M. conjunctivae			1	3
	M. crocodyli	1		1	1
	M. fermentans			2	2
	M. gallinaceum	2[h]	2[h]	2	
	M. hyopneumoniae				3
	M. hyorhinis				1
	M. mobile	1	1	1	1
	M. pulmonis	1		1	1
	M. synoviae	1	3	1	3
	M. arginini				
	M. arthritidis				
	M. bovoculi				
	M. canadense				
	M. dispar				
	M. flocculare				
	M. hominis				
Spiroplasma group	*M. capricolum* subsp. *capricolum*			1	1
	M. capricolum subsp. *capripneumoniae*			2	2
	M. leachii			1	1
	M. mycoides subsp. *mycoides*			1	1
	M. mycoides subsp. *capri*			1	1
	M. putrefaciens			2	2
	M. yeatsii			2	2
Pneumoniae group	*M. gallisepticum*				1
	M. genitalium				1
	M. penetrans	1		1	1
	M. pneumoniae				1
	M. haemocanis				
	M. ovis				
	M. parvum				
	M. suis				
	M. wenyonii				

*CoaBC is annotated as separate PPCS and PPCDC in the databases
[a]Based on 16S rRNA sequence data
[h]Hypothetical protein (as annotated on NCBI)
1 = according to both KEGG and SEED viewer
2 = according to KEGG only
3 = according to SEED viewer only
M. species = *Mycoplasma* species with no pathway enzymes according to KEGG

Table 3.2 The coenzyme A biosynthetic pathway enzyme-encoding genes in the 62 investigated *Mycoplasma* species, along with the hosts and energy sources of the corresponding *Mycoplasma* species

Phylogenetic group[a]	Mycoplasma species	Representative host	Energy source[b]	CoA biosynthetic pathway enzyme			
				PanK	CoaBC	PPAT	DPCK
Hominis group	*M. anatis**	Ducks	G	1	2	3	4
	*M. arginini**	Mammals	R	1	2	3	4
	*M. columborale**	Pigeons	G	1	2	3	4
	*M. cricetuli**	Hamsters	G	1	2	3	4[h]
	M. mobile^	Tench	G, R	1	2	3[p]	4
	*M. sturni**	Songbirds	G	1	2	3	4[h]
	M. synoviae^	Galliforms	G	1	2	3[p]	4[h]
	M. gallinaceum^	Galliforms	G	1[h]	2[h]	3	
	*M. alligatoris**	Alligators	G	1		3	4[h]
	*M. buteonis**	Raptors	G	1		3	4[h]
	M. crocodyli^	Crocodiles	G	1		3	4
	*M. molare**	Dogs	G	1		3	4
	M. pulmonis^	Mice	G	1		3	4
	M. agalactiae^	Goats	G			3	4
	M. bovigenitalium^	Cattle	OH, OA			3	4
	M. bovis^	Cattle	OH, OA			3	4
	M. californicum^	Cattle	OH, OA			3	4
	*M. canis**	Dogs	G			3	4[h]
	*M. collis**	Rodents	G			3	4
	*M. columbinum**	Pigeons	R			3	4
	*M. felifaucium**	Pumas	R			3	4
	*M. felis**	Cats	G			3	4
	M. fermentans^	Humans	R, G			3	4
	*M. gallinarum**	Galliforms	R			3	4
	*M. iners**	Galliforms	R			3	4
	*M. leonicaptivi**	Lions	G			3	4

Phylogenetic group[a]	Mycoplasma species	Representative host	Energy source[b]	CoA biosynthetic pathway enzyme			
				PanK	CoaBC	PPAT	DPCK
Hominis group	*M. lipofaciens**	Galliforms	G, R			3	4
	*M. opalescens**	Dogs	R			3	4
	*M. primatum**	Monkeys	R			3	4
	*M. simbae**	Lions	R			3	4
	M. conjunctivae^	Goats	G			3	5
	M. bovoculi^	Cattle	G				5
	M. dispar^	Cattle	G				5
	M. flocculare^	Pigs	U				5[h]
	M. hyopneumoniae^	Pigs	G				5
	*M. ovipneumoniae**	Sheep	G				5
	M. hyorhinis^	Pigs	G				4
	M. sp. Ms02*	Ostrich	U				4
	M. arthritidis^	Rats	R				
	M. canadense^	Cattle	R				
	M. hominis^	Humans	R				
Spiroplasma group	*M. capricolum* subsp. *capricolum*^	Goats	G			3	4
	M. capricolum subsp. *capripneumoniae*^	Goats	G			3	4
	M. leachii^	Cattle	G			3	4
	M. mycoides subsp. *capri* LC^	Goats	G			3	4
	M. mycoides subsp. *mycoides* SC^	Cattle	G			3	4
	M. putrefaciens^	Goats	G			3	4
	M. yeatsii^	Goats	G			3	4
Pneumoniae group	*M. iowae**	Turkeys	G, R	1	2s[c]	3	4
	*M. testudinis**	Tortoises	G	1	2[h]	3	4
	*M. alvi**	Cattle	G, R	1		3	4
	M. penetrans^	Humans	G, R	1		3	4
	*M. pirum**	Humans	G	1		3	4
	M. gallisepticum^	Galliforms	G				4

Phylogenetic group[a]	Mycoplasma species	Representative host	Energy source[b]	CoA biosynthetic pathway enzyme			
				PanK	CoaBC	PPAT	DPCK
Pneumoniae group	*M. genitalium*^	Humans	G				4
	*M. imitans**	Ducks, geese	G				4
	M. pneumoniae^	Humans	G				4
	M. haemocanis^	Dogs	U				
	M. ovis^	Sheep	U				
	M. parvum^	Pigs	U				
	M. suis^	Pigs	U				
	M. wenyonii^	Cattle	U				

^Complete genome (deposited in NCBI)
*Contig assembled genome (deposited in NCBI)
[a]Based on 16S rRNA sequence data
[b]Abbreviations: G, glucose; OH, alcohols; OA, organic acids; R, arginine; U, undefined
[c]Separate PPCS and PPCDC enzymes (no bifunctional protein)
[h]Hypothetical protein (as annotated)
[p]Putative protein (as annotated)
1 = PanK type III
2 = CoaBC
3 = PPAT
4 = DPCK
5 = HAD-DPCK
M. species = *Mycoplasma* species with no pathway enzymes identified

3.4.2 Identification of CoA biosynthesis enzyme-encoding gene homologues

At the time of this investigation, no pathway enzymes could be identified in eight of the 62 *Mycoplasma* species genomes available on NCBI. Furthermore, there were ten hypothetical- and two putative CoA biosynthetic enzyme-encoding gene homologues identified (Table 3.2). The two putative gene homologues, which were identified in *M. synoviae* and *M. mobile*, are believed to encode PPAT. Of the ten hypothetical gene homologues, one encodes a type III PanK enzyme (identified in *M. gallinaceum*) and two encode CoaBC enzymes (*M. gallinaceum and Mycoplasma testudinis*). The remaining seven encode DPCK enzymes, of which one was identified as a HAD-DPCK enzyme (*Mycoplasma flocculare*).

3.4.3 Analysis of genome locations of enzyme-encoding genes and the association between number of genes and energy source used by species or its host

The genomic location of each currently annotated enzyme-encoding gene and newly identified homologues, within the corresponding mycoplasma genome, was determined (Supplementary Table 1.2). At the time of the study, many of the genomes of the investigated *Mycoplasma* species were only available in contig sets; not complete genome sequences. This made it difficult to identify patterns concerning the genomic locations of the CoA biosynthetic enzyme-encoding genes, with respect to each other. Nevertheless, of the 62 investigated *Mycoplasma* species, there were 35 with complete genomes (Table 3.2). All the species in the spiroplasma group, which fall within these 35 species, displayed a distance in the range of 122-224 kb between the DPCK gene location and that of the PPAT gene. In addition, the only three species with complete genomes and an identified CoaBC gene revealed that the PanK and CoaBC genes have overlapping coding regions. A similar trend of overlapping coding regions concerning these two genes is observed in the other species with genomes consisting of contigs, as these genes are located on the same contig. However, even though the PPCDC and PPCS encoding genes of *M. iowae* are located on one contig (separated by 5 bp), the PanK encoding gene is located on a different contig.

There seems to be no correlation between the general energy source of the *Mycoplasma* species and their number of identified CoA biosynthetic pathway enzymes. There also seems to be no correlation between the host of the species and the number of identified CoA biosynthetic pathway enzymes.

3.4.4 Bioinformatics analysis of annotated and newly identified CoA biosynthesis enzyme-encoding gene homologues

3.4.4.1 PanK type III analyses

The CDD searches of the identified type III PanK protein sequences (Supplementary Table 1.3) confirmed the identity of all of the proteins (defined as PanK type III), except the one present in *Mycoplasma anatis* (for which no conserved domain could be identified). CDD could also predict that all of these proteins belong to the ASKHA superfamily (cl17037), which is as expected for type III PanKs. The InterPro searches (Supplementary Table 1.7) supported the CDD results, as it could predict that all of the proteins, including the *M. anatis* protein, belong to the Interpro family IPR004619 (Type III pantothenate kinase), along with the predicted molecular function gene ontology (GO) term, GO:0004594 (pantothenate kinase activity). Additionally, the region of identification within the hypothetical PanK sequence of *M. gallinaceum* was similar to the annotated type III PanKs in both the CDD and InterPro searches.

The typical length of the annotated type III PanK amino acid sequences is between 220 aa and 275 aa. The hypothetical PanK protein sequence identified in *M. gallinaceum* falls within this range with a length of 248 aa, which further supports the identity of this hypothetical protein as a type III PanK.

All four of the motifs identified by the MEME searches (Supplementary Table 1.12) overlapped with the identified regions predicted by CDD and InterPro. However, the protein sequence of two species (*M. anatis* and *Mycoplasma buteonis*) was found to differ from the rest, as no Motif 2 and Motif 4 were identified in these sequences. Motif 4 was also shown to be absent in the protein sequences of four other species, namely *Mycoplasma mobile*, *Mycoplasma molare*, *Mycoplasma pulmonis*, and *M. testudinis*.

3.4.4.2 CoaBC analyses

The CDD searches of the identified CoaBC protein sequences of all hypothetical and annotated proteins, as well as the separate PPCS and PPCDC protein sequences present in *M. iowae* (Supplementary Table 1.4), confirmed their identity as CoaBC proteins. The proteins were all predicted to belong to the DNA/pantothenate metabolism flavoprotein (DFP) superfamily (cl27193), with a large region (almost the entire sequence length) of identification. Another superfamily predicted by CDD was the Flavoprotein superfamily (cl19190). However, this region of identification was observably smaller (about half of the total sequence length) and located at the N-terminal end of all the protein sequences apart from the PPCS protein of *M. iowae*. Moreover, the separate PPCS and PPCDC sequences,

each had a significant hit defined as streptococcal PPCS and streptococcal PPCDC, respectively.

The InterPro searches (Supplementary Table 1.8) identified the InterPro family IPR005252 signature (Coenzyme A biosynthesis bifunctional protein, CoaBC) in most of the CoaBC proteins, including the hypothetical protein of *M. gallinaceum*. For these protein sequences, InterPro also predicted the biological process GO term GO:0015937 (CoA biosynthetic process) and molecular function GO terms, GO:0004632 (PPCS activity) and GO:0004633 (PPCDC activity). Although no InterPro family IPR005252 signature was identified in the CoaBC protein sequence of *M. mobile* and the hypothetical CoaBC protein sequence of *M. testudinis*, the InterPro domain signatures, IPR003382 (Flavoprotein) and IPR007085 (DFP, C-terminal), were identified, which is also present in all of the other protein sequences. These respective InterPro domain signatures were also identified in the *M. iowae* PPCDC protein sequence (IPR003382) and the *M. iowae* PPCS protein sequence (IPR007085). Furthermore, InterPro also predicted a signal region at the N-terminal for most of the CoaBC protein sequences, with the exception of *M. gallinaceum*, *M. testudinis* and *M. mobile*. Similarly, a “Prokaryotic membrane lipoprotein lipid attachment site” was identified at the N-terminal in the *M. iowae* PPCDC protein sequence.

The average length of the annotated CoaBC amino acid sequences ranges between 360 aa and 410 aa. The two hypothetical CoaBC proteins are within this range with 385 aa and 380 aa for *M. gallinaceum* and *M. testudinis*, respectively. This provides additional support for the identity confirmation of these two hypotheticals as CoaBC proteins.

The regions of the MEME identified motifs (Supplementary Table 1.13) corresponded with the identified regions predicted by CDD and InterPro. Two of the motifs (Motif 1 and Motif 2) were located in the PPCDC domain of the bifunctional protein, whilst Motif 3 and Motif 4 were found in the PPCS domain. Similarly, of the four motifs, only Motif 1 and Motif 2 could be identified in the PPCDC protein sequence of *M. iowae*, which was as expected since PPCDC is the N-terminal domain of the CoaBC protein. Conversely, in its PPCS protein sequence, only Motif 3 could be identified. Moreover, the protein sequences of *M. mobile* and *M. testudinis* were the only two CoaBC protein sequences with no identified Motif 4.

3.4.4.3 PPAT analyses

The CDD results of the identified PPAT protein sequences (Supplementary Table 1.5) confirmed the identity of all of the hypothetical- and annotated proteins (defined as PPAT). However, the region of this identification in the *M. mobile* putative protein sequence was noticeably shorter (66% of the total sequence length) compared to that observed in the other

organisms. Additionally, CDD predicted that all these proteins belong to the nucleotidyltransferase superfamily (cl00015). Similar results were obtained from the InterPro searches (Supplementary Table 1.9), where the InterPro family signature, IPR001980 (Phosphopantetheine adenylyltransferase), was identified in all of the protein sequences. As observed in the CDD results, the identified region within the *M. mobile* putative protein sequence was not similar to the other organisms. Furthermore, the InterPro predicted GO term for the biological process and molecular function was GO:0015937 (CoA biosynthetic process) and GO:0004595 (PPAT activity), respectively.

The typical length of the annotated PPAT protein sequences ranged between 135 aa and 175 aa. Both putative PPAT proteins fall within this range with respective protein lengths of 145 aa and 148 aa for *M. mobile* and *M. synoviae*.

Both the CDD and InterPro identified regions in the protein sequences overlap with the four MEME identified motifs (Supplementary Table 1.14). Yet again, the protein sequence of *M. mobile* was the outlier, in which only two motifs (Motif 1 and Motif 2) could be identified. Also, no Motif 3 could be identified within the protein sequences of *Mycoplasma alvi*, *Mycoplasma penetrans*, *Mycoplasma pirum*, *M. iowae*, and *M. testudinis*.

3.4.4.4 DPCK analyses

The identity of all the DPCK protein sequences could be confirmed using CDD searches (defined as DPCK, Supplementary Table 1.6). All of the sequences, including the HAD-DPCK protein sequences, were predicted to belong to the nucleoside/nucleotide kinase (NK) superfamily (cl17190) and the CoaE superfamily (cl28605). In addition, CDD could predict that the six HAD-DPCK protein sequences, which included the hypothetical protein of *M. flocculare*, belong to the HAD-like superfamily (cl21460) and the Hydrolase 3 superfamily (cl26787). This HAD-like protein domain was further predicted to belong to the Cof subfamily, which falls within the Class-IIB subfamily of the HAD-like superfamily. A specific hit, providing only a general function prediction, identified this protein sequence as Cof, referring to hydroxymethylpyrimidine pyrophosphatase and other HAD family phosphatases.

In the InterPro searches (Supplementary Table 1.10), the InterPro family signature, IPR001977 (Dephospho-CoA kinase), was identified in all of the annotated protein sequences. The corresponding GO terms predicted for these proteins were: GO:0015937 (CoA biosynthetic process), referring to the biological process; along with GO:0004140 (DPCK activity) and GO:0005524 (ATP binding), referring to the molecular function. These signatures could, however, not be identified in the seven hypothetical protein sequences that belong to *Mycoplasma canis*, *Mycoplasma cricetuli*, *Mycoplasma sturni, Mycoplasmas*

alligatoris, M. flocculare, M. buteonis, and *M. synoviae*. The only InterPro signature that could be identified in these protein sequences was the homologous superfamily signature, IPR027417 (P-loop containing nucleoside triphosphate hydrolase), which was identified in the protein sequences of all the species, including annotated protein sequences. However, the CDD predictions of the protein sequences of the species in question suggest that they are indeed DPCK. With regards to the HAD-like domain of the HAD-DPCK proteins, which included the hypothetical protein of *M. flocculare*, InterPro could only identify the family signature, IPR006379 (HAD-superfamily hydrolase, subfamily IIB).

The average length of the annotated DPCK amino acid sequences is between 165 aa and 205 aa, where the average annotated HAD-DPCK amino acid sequence length is ±445 aa. The six hypothetical DPCK proteins of *M. canis* (189 aa), *M. cricetuli* (191 aa), *M. sturni* (187 aa), *M. alligatoris* (185 aa), *M. buteonis* (189 aa) and *M. synoviae* (168 aa) fall within the average DPCK protein range. Similarly, the hypothetical HAD-DPCK protein of *M. flocculare* (447 aa) is also within the range of the average HAD-DPCK protein sequence length. This offers additional support for the confirmation of the identity of these hypotheticals.

All four of the MEME predicted motifs (Supplementary Table 1.15) overlapped with the regions identified by the CDD and InterPro searches. The first three motifs were present in all of the sequences, with a noticeable difference in the identification regions of the HAD-DPCK protein sequences. However, if the HAD-domain was to be removed, these motif regions would match that of the motif regions predicted in the protein sequences of the other species. Moreover, MEME could not identify Motif 4 within 17 protein sequences (of which none were hypothetical proteins).

3.4.5 Multiple sequence alignments and phylogenetic analyses

The phylogeny generated based on the 16S rRNA gene sequences of the *Mycoplasma* species (Figure 3.1) revealed a clear distinction between three groups i.e. spiroplasma, pneumoniae, and hominis. These three group divisions were resolved and well supported with bootstrap values above 75%. Additionally, there were identifiable clades within these groups, which were also well supported with high bootstrap values. Some of the clades seem to show consistency relating to the number of identified CoA biosynthetic pathway enzymes in the species, whereas other clades seem to be inconsistent in this regard.

A phylogenetic analysis of the amino acid sequences for each of the CoA biosynthetic pathway enzymes was subsequently performed. All of the investigated mycoplasmas are, however, not presented in each of the four phylogenies since not all the species possess all four pathway enzymes.

The MSAs of the respective proteins used for phylogenetic analysis consisted of amino acid sequences from: (i) 18 *Mycoplasma* species for the PanK type III enzyme (Supplementary Figure 1.1); (ii) 9 *Mycoplasma* species for the CoaBC enzyme with *M. iowae* excluded (Supplementary Figure 1.2); (iii) 10 *Mycoplasma* species with an additional CoaBC enzyme with a concatenated sequence of the PPCDC and PPCS sequences of *M. iowae* (producing an artificial CoaBC sequence, Supplementary Figure 1.3); (iv) 42 *Mycoplasma* species for the PPAT enzyme (Supplementary Figure 1.4); and (v) 53 *Mycoplasma* species for the DPCK enzyme (Supplementary Figure 1.5).

The relevant amino acid sequences of closely related species, namely *C. innocuum*, *L. fermentum*, *S. pneumoniae* and *B. coahuilensis*, were included in each of these phylogenetic analyses (Table 3.3). However, no type III pantothenate kinase (PanK) protein could be identified in *L. fermentum,* as it only possesses a type I PanK; nor could a bifunctional CoaBC protein be identified in *S. pneumoniae*, as it possesses separate PPCS and PPCDC proteins.

Table 3.3 The related organisms used in the phylogenetic analyses and the NCBI accession numbers of their respective CoA biosynthetic proteins

Organism	**NCBI accession number**			
	PanK type III	**CoaBC**	**PPAT**	**DPCK**
B. coahuilensis m2-6	WP_010169518.1	WP_082688052.1	WP_059282518.1	WP_059283260.1
C. innocuum I46	ASU20925.1	ASU20926.1	ASU18107.1	ASU18339.1
L. fermentum IFO 3956	-	WP_012391452.1	BAG26942.1	BAG27651.1
S. pneumoniae N	CKG77743.1	-	CKE32357.1	CKG66021.1

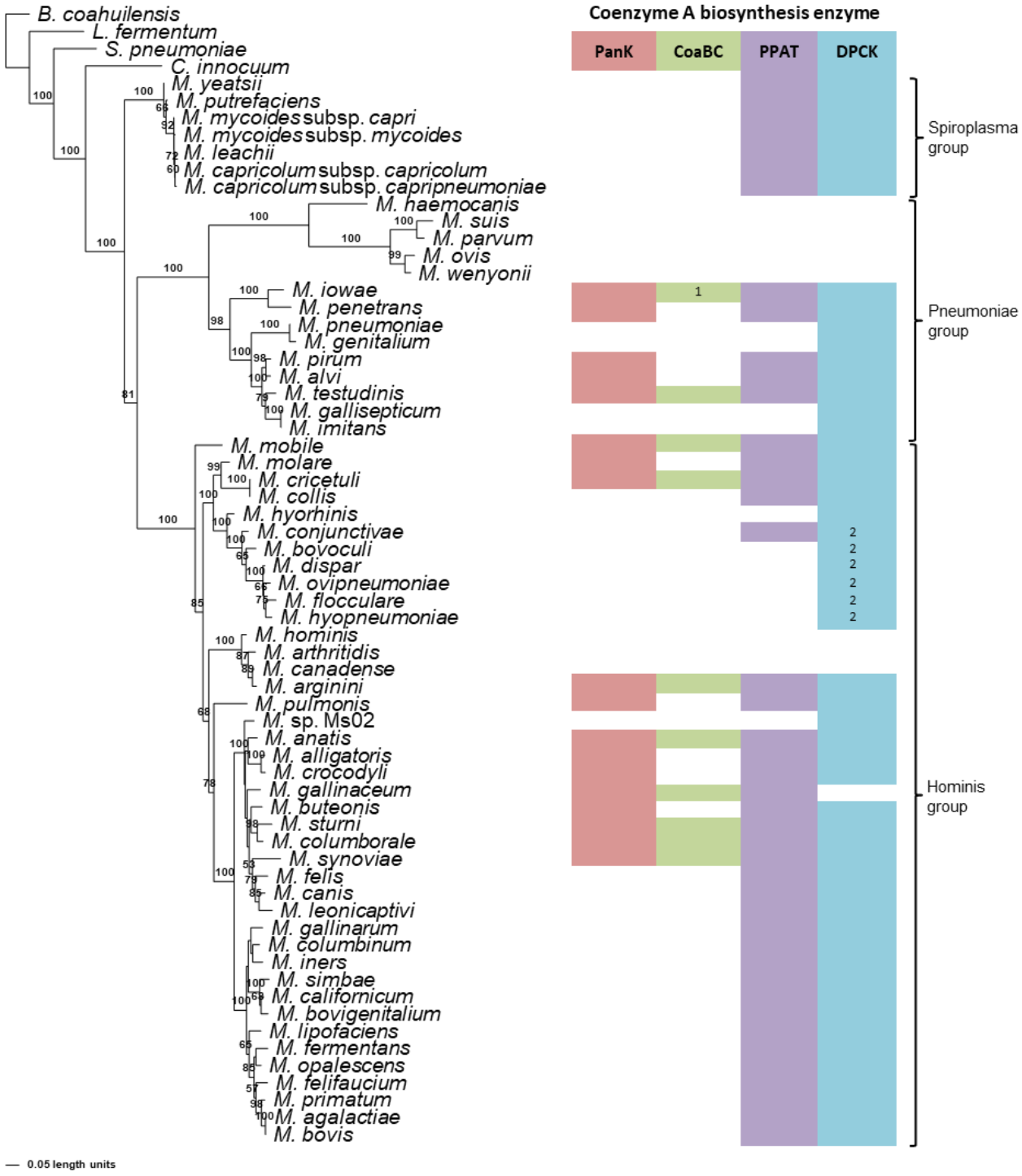


Figure 3.1 The maximum likelihood phylogeny based on 16S rRNA sequences. Bootstrap values (≥50%) are indicated above lines. The identified coenzyme A biosynthetic pathway enzymes of each *Mycoplasma* species are indicated with a colour chart: PanK (red), CoaBC (green), PPAT (purple), DPCK (blue). Notations: 1 (green) – separate PPCS and PPCDC enzymes; 2 (blue) – HAD-DPCK protein.

The phylogeny based on the PanK type III protein sequences (Figure 3.2) revealed two well-supported clades. The distribution of species within clades did not agree with that of the 16S rRNA sequence-based phylogeny. There is also no correlation between the species in a clade and the number of identified CoA biosynthetic pathway enzymes within these species. Similarly, at first glance, this seemed to be the case for the MEME motif predictions as well. However, after examining the MSA of the PanK type III protein sequences (Supplementary Figure 1.1) at the MEME-predicted Motif 2 site, which was absent in *M. anatis* and *M. buteonis*, it was revealed that the sequences of these species contain insertions in this region, at positions 184-186 and 179, respectively. The MSA also revealed similar mutation incidents at the predicted Motif 4 site in which the protein sequences with a reportedly absent motif had deletions within the predicted region. This predicted Motif 4 in the protein sequences of *M. pulmonis*, *M. molare*, *M. mobile* and *M. testudinis* seem to have similar sequences and deletions with respect to each other (deletion located at positions 236-238 in the MSA), but they differ compared to that of *M. buteonis*. *M. anatis*, however, had an entirely different sequence at this predicted motif region.

Therefore, there does seem to be some correlation between the MEME motif predictions of the protein sequences that belong to *M. pulmonis*, *M. molare*, *M. mobile* and *M. testudinis* since these four are located within the same clade in the PanK type III sequence-based phylogeny and their protein sequences were all identified to have an absent Motif 4. The same, however, cannot be said for the MEME motif predictions of the *M. anatis* and *M. buteonis* protein sequences, as they were both predicted to have an absent Motif 2 and Motif 4, yet they are not in the same clade according to the PanK type III-based phylogeny; in fact, they are quite distant.

Similar to the PanK type III sequence-based phylogeny, both phylogenies based on the CoaBC protein sequences (Figure 3.3-3.4) showed no resemblance to the 16S rRNA sequence-based phylogeny concerning species distribution. The absent MEME-predicted Motif 4 in the protein sequences of *M. mobile* and *M. testudinis* correspond to their position in the CoaBC sequence-based phylogeny, as they are situated in an isolated clade. The MSA of the CoaBC protein sequences (Supplementary Figure 1.2) further confirmed the relationship of these two protein sequences, as a similar large deletion (positions 256-275) can be observed in both sequences. Additionally, comparison of the two CoaBC sequence-based phylogenies revealed that the concatenated sequence of the separate *M. iowae* PPCS and PPCDC protein sequences was retrieved in a basal position (well supported by the Bootstrap analysis) in the relevant phylogeny (Figure 3.4), as expected.

As with 16S rRNA sequence-based phylogeny, a distinctive and well-supported clade for the spiroplasma group could be retrieved in the phylogeny based on the PPAT protein sequences (Figure 3.5). The pneumoniae- and hominis groups could also be observed but displayed poor clade support with no prominent clades. The species with no MEME-predicted Motif 3 are all species that fall within the pneumoniae group according to their 16S rRNA sequence-based phylogenetic relationship, except *M. mobile*. The MSA of the PPAT protein sequences (Supplementary Figure 1.4) revealed that these sequences, including that of *M. mobile*, are similar to each other regarding the deletion within the absent MEME-predicted motif region.

Similar to the PPAT sequence-based phylogeny, the phylogeny based on the DPCK protein sequences (Figure 3.6) could retrieve a distinctive and well-supported clade for the spiroplasma group, but also displayed poor clade support in the pneumoniae- and hominis groups. However, distinctive clades were retrieved in the DPCK protein sequence-based phylogeny (albeit poorly supported), which resemble those observed in the phylogeny based on the 16S rRNA sequences. Moreover, the *Mycoplasma* species, in which MEME-predicted Motif 4 was predicted to be absent, were all species from either the spiroplasma group or the pneumoniae group. After examining the MSA of the DPCK sequences (Supplementary Figure 1.5), it was clear that the species with an absent Motif 4 had noticeable differences in their DPCK protein sequences with respect to the other species.

There was, however, one recurring difference between the 16S rRNA sequence-based phylogeny and all the protein sequence-based phylogenies, and that is the location of *Mycoplasma arginini*.

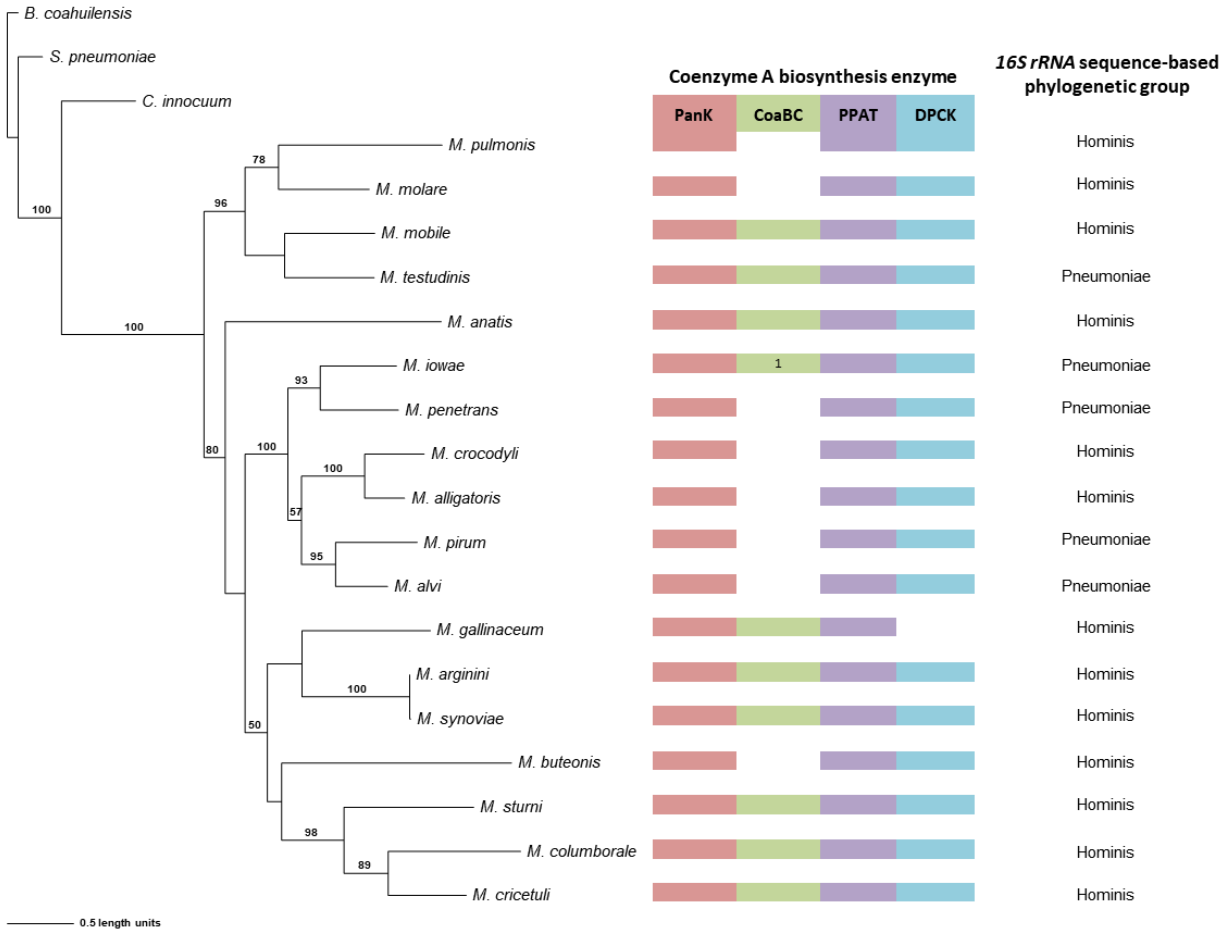


Figure 3.2 The maximum likelihood phylogeny based on the PanK type III protein sequences. Bootstrap values (≥50%) are indicated above lines. Grouping of the species according to the 16S rRNA sequence-based phylogeny, indicated on the right. The identified coenzyme A biosynthetic pathway enzymes of each *Mycoplasma* species are indicated with a colour chart: PanK (red), CoaBC (green), PPAT (purple), DPCK (blue). Notation: 1 (green) –separate PPCS and PPCDC enzymes.

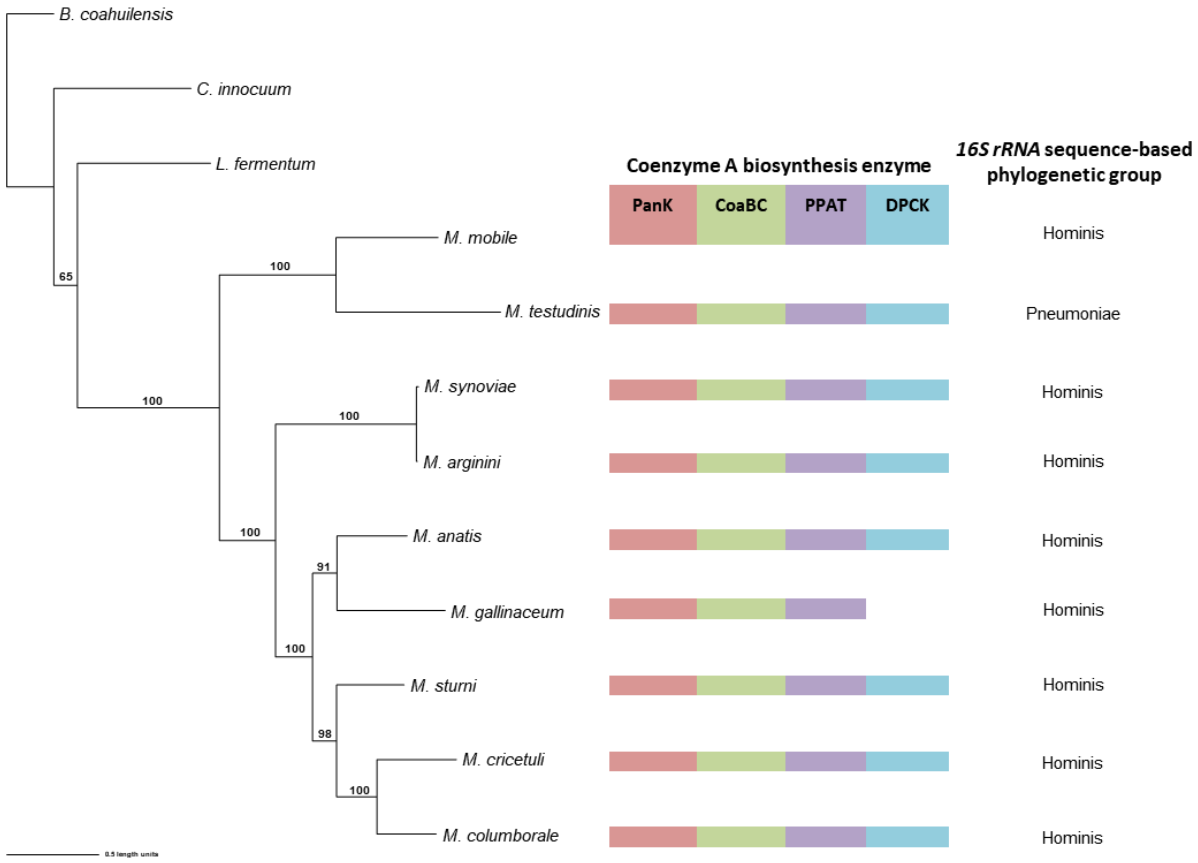


Figure 3.3 The maximum likelihood phylogeny based on the CoaBC protein sequences, excluding *M. iowae*. Bootstrap values (≥50%) are indicated above lines. Grouping of the species according to the 16S rRNA sequence-based phylogeny, indicated on the right. The identified coenzyme A biosynthetic pathway enzymes of each *Mycoplasma* species are indicated with a colour chart: PanK (red), CoaBC (green), PPAT (purple), DPCK (blue). Notation: 1 (green) –separate PPCS and PPCDC enzymes.

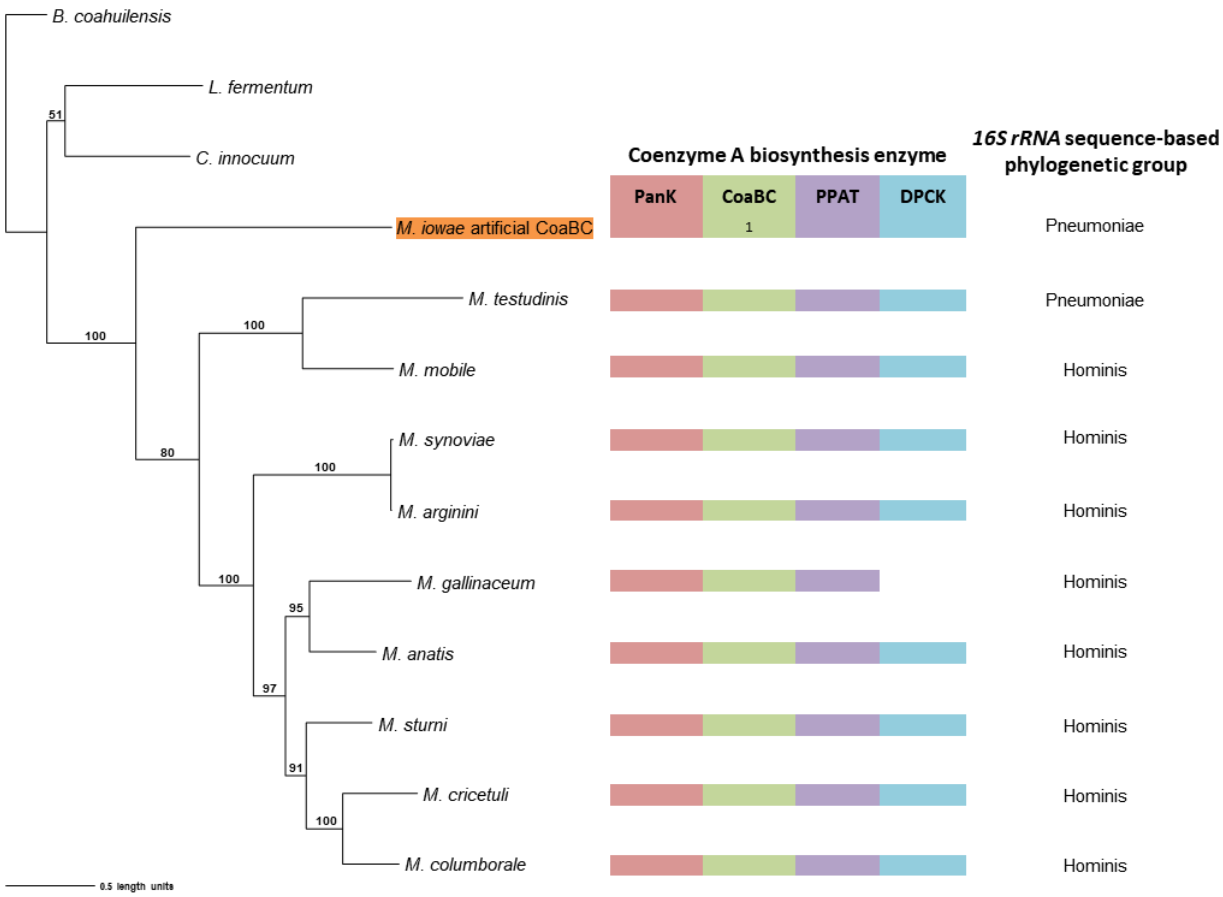


Figure 3.4 The maximum likelihood phylogeny based on the CoaBC protein sequences, including *M. iowae*. Bootstrap values (≥50%) are indicated above lines. Grouping of the species according to the 16S rRNA sequence-based phylogeny, indicated on the right. The identified coenzyme A biosynthetic pathway enzymes of each *Mycoplasma* species are indicated with a colour chart: PanK (red), CoaBC (green), PPAT (purple), DPCK (blue). Notation: 1 (green) –separate PPCS and PPCDC enzymes. The orange highlighted species is the concatenated PPCDC and PPCS protein sequence.

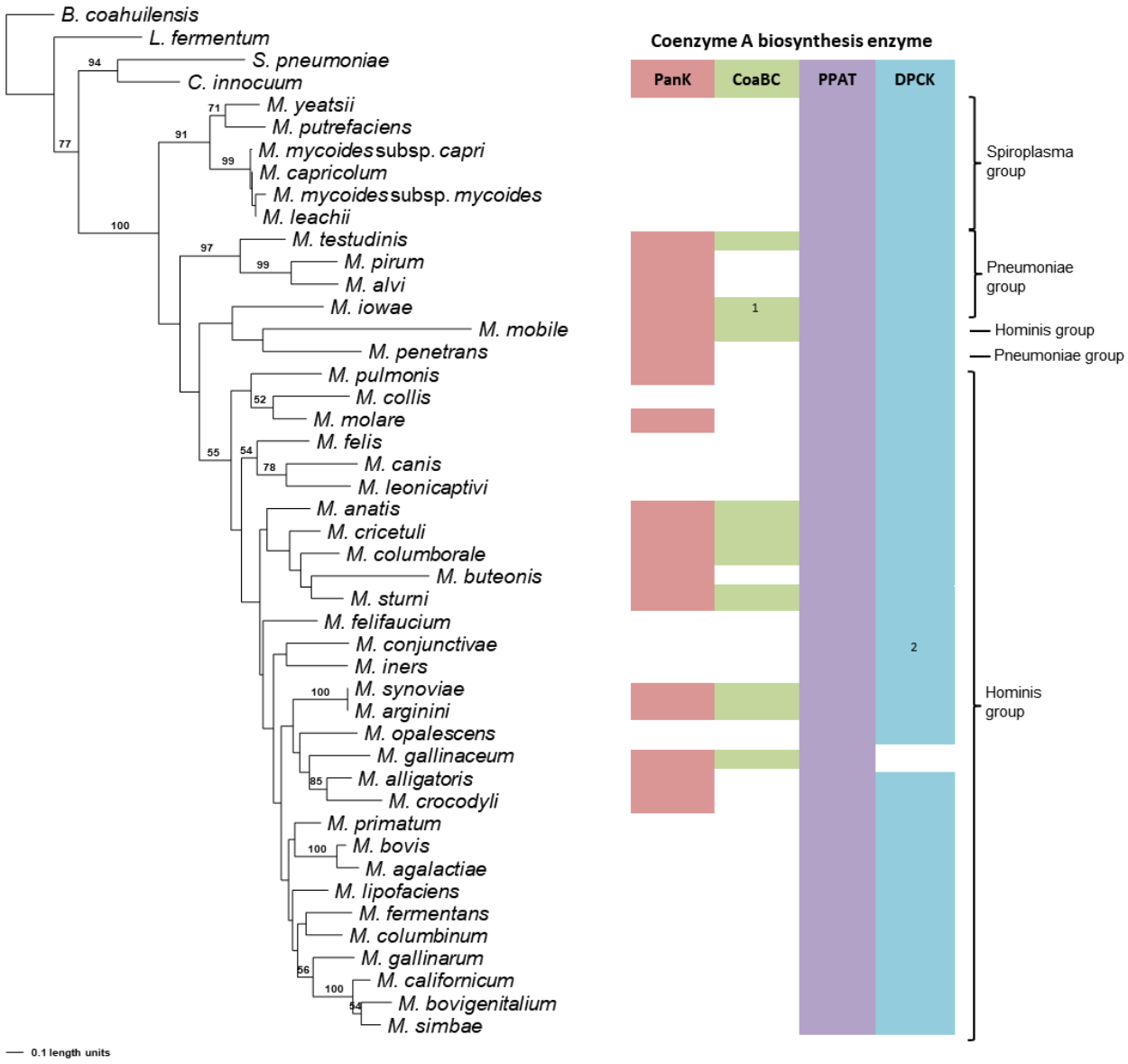


Figure 3.5 The maximum likelihood phylogeny based on the PPAT protein sequences. Bootstrap values (≥50%) are indicated above lines. Suggestive groupings, which resemble that of the 16S rRNA sequence-based phylogeny, are indicated by the brackets (right), with *M. mobile* and *M. penetans* as exceptions. The identified coenzyme A biosynthetic pathway enzymes of each *Mycoplasma* species are indicated with a colour chart: PanK (red), CoaBC (green), PPAT (purple), DPCK (blue). Notations: 1 (green) – separate PPCS and PPCDC enzymes; 2 (blue) – HAD-DPCK protein.

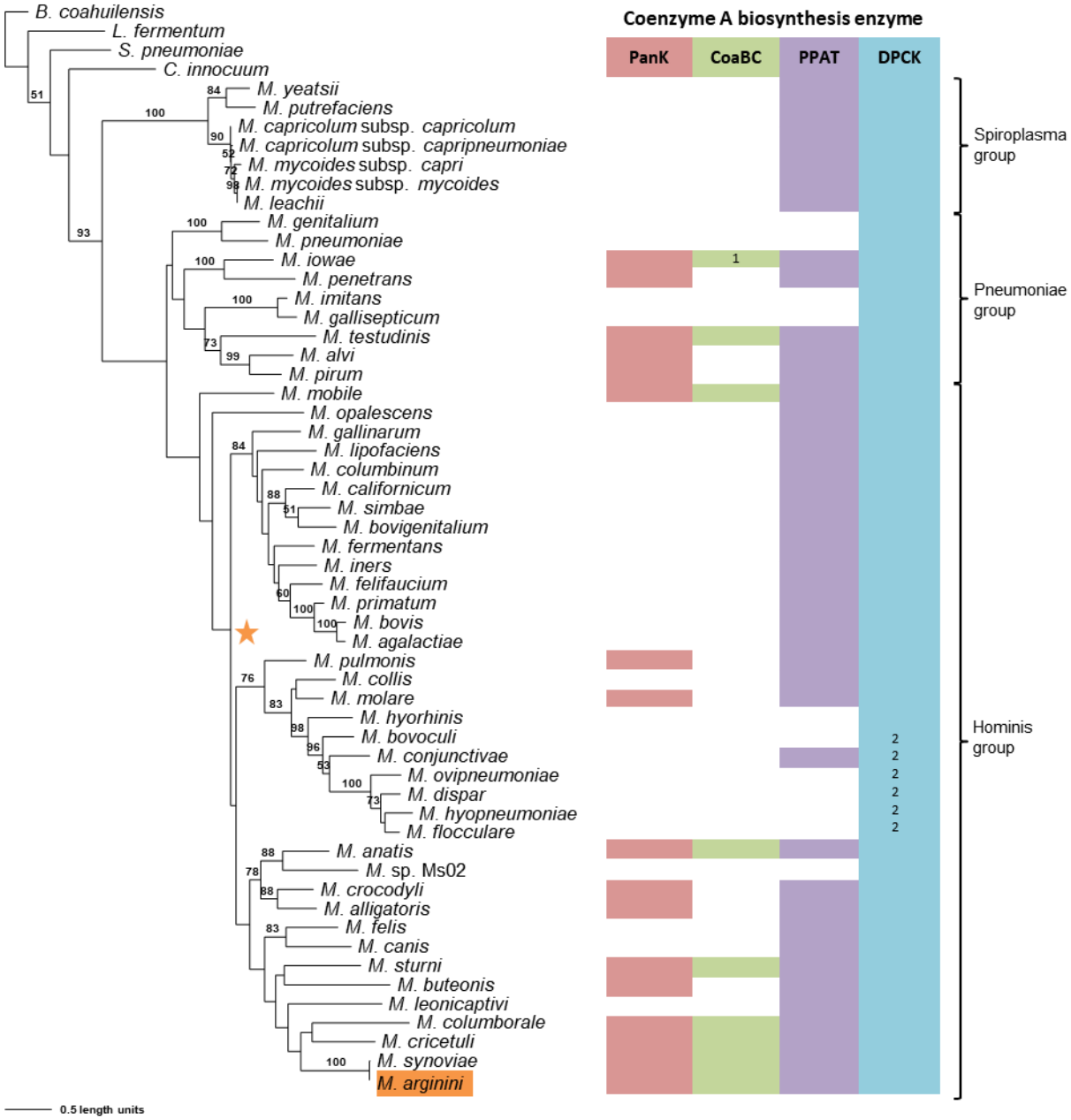


Figure 3.6 The maximum likelihood phylogeny based on the DPCK protein sequences. Bootstrap values (≥50%) are indicated above lines. Suggestive groupings of the three prominent clades, which resemble that of the 16S rRNA sequence-based phylogeny, are indicated by the brackets (right). The identified coenzyme A biosynthetic pathway enzymes of each *Mycoplasma* species are indicated with a colour chart: PanK (red), CoaBC (green), PPAT (purple), DPCK (blue). Notations: 1 (green) – separate PPCS and PPCDC enzymes; 2 (blue) – HAD-DPCK protein. The orange highlighted species is the sub-clade species variation from the 16S rRNA sequence-based phylogeny. The orange star represents the location of the species in the 16S rRNA sequence-based phylogeny for *M. arginini*.

3.5 Discussion

The phylogenetic analysis of the 16S rRNA sequences produced a phylogeny (Figure 3.1) that reflects the phylogenetic relationship of the investigated mycoplasmas, which is very similar to what is reported in the literature [28–30, 45]. Moreover, the prominent clades that could be observed in the retrieved phylogeny match the distinctive clusters as reported by Brown [26].

The number of *Mycoplasma* species with annotated genes available on NCBI and identified using BLAST searches was much more than that reported on both the KEGG pathway and SEED viewer subsystems databases. This might be due to the number of complete genomes available for *Mycoplasmas* species. Another reason that might explain the observed inconsistencies is the fact that these databases are extremely large and although the KEGG pathway and SEED viewer subsystems databases are updated continuously [187, 190], it can be very difficult to keep each aspect of the database up to date.

The identity of each hypothetical protein sequence was determined based on motif, domain and family relationships and subsequently compared to that of the currently annotated genes. Results revealed that all of the identified PanK protein sequences were type III PanK enzymes. Furthermore, these proteins were predicted to belong to the ASKHA superfamily, which is in agreement with the literature on the type III PanKs [103, 127, 132]. In addition, there were segments within the MEME-predicted motifs that are in agreement with known motifs of the ASKHA superfamily, as well as type III PanK proteins [103, 132, 207, 208]. However, there was one motif (the Pan Motif) that seem to have undergone a substitution mutation in all of the identified mycoplasma type III PanK proteins (Figure 3.7). The Pan Motif normally bears a conserved residue sequence of *h*G*h*DR (where *h* = hydrophobic residue) [207]. Yet, in the identified mycoplasma protein sequences, the arginine residue is replaced by a leucine or isoleucine residue. This arginine residue plays a prominent rule in the active site of the PanK enzyme since it is involved in an interaction with pantothenate. Therefore, the substitution of this charged residue with an uncharged hydrophobic residue will most likely have major consequences with regards to substrate binding, specifically pantothenate, and might allow PantSH to be used as a substrate. However, this should be tested experimentally.

Additionally, the MEME results indicated that some of the sequences lack certain motifs. The MEME algorithm that was used for these searches only detects ungapped motifs. Therefore, if a protein sequence contains mutations, such as insertions or deletions, within the predicted motif region, the motif in question will subsequently be regarded absent. This can

be seen by examining the MSA of the relevant enzyme's protein sequences. An example of this can be seen at the predicted Motif 2 site in the MSA of the type III PanK protein sequences, where the motif was regarded as absent in *M. anatis* and *M. buteonis* due to insertions at positions 184-186 for *M. anatis* and 179 for *M. buteonis* (Figure 3.7). These two species also lacked a MEME-predicted Motif 4, but the reasons for this were different in the respective protein sequences of the two species. This might, thus, explain why these two species are quite distantly located from each other on the PanK type III-based phylogeny, even though they have the same absent MEME-predicted motifs. Conversely, the deletions observed at the predicted Motif 4 in the protein sequences of *M. pulmonis*, *M. molare*, *M. mobile* and *M. testudinis* might suggest a functional relationship between the PanKs of these species, as supported by their clade grouping in the PanK type III-based phylogeny.

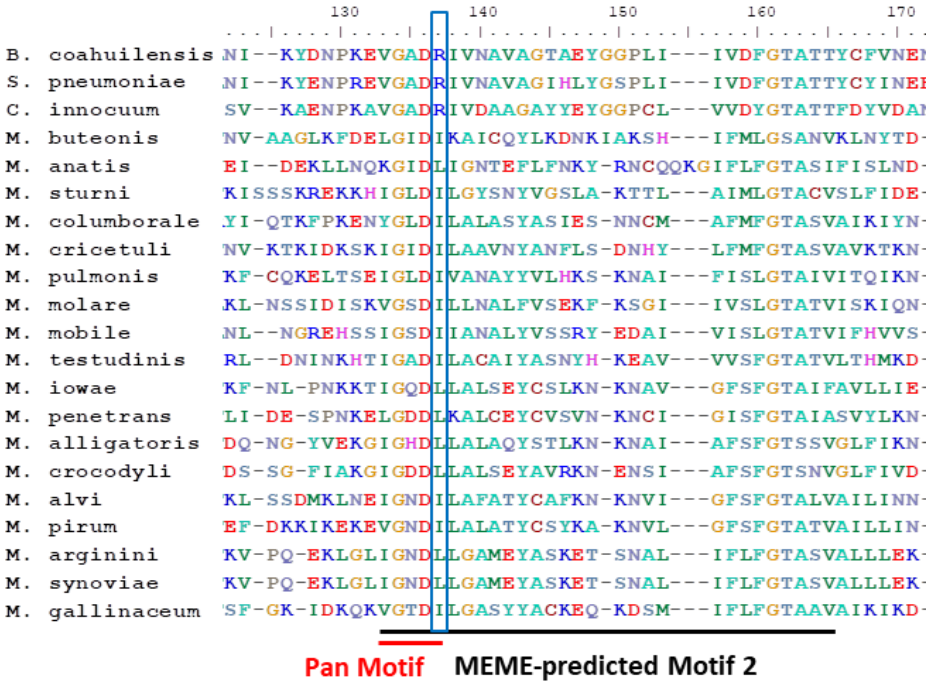


Figure 3.7 A section of the multiple sequence alignment of the type III PanK protein sequences. The Pan Motif (red) and MEME-predicted Motif 2 (black) are underlined. Insertions are indicated by dashes (-). The mycoplasma Pan Motif mutation site is indicated by the blue box.

The results obtained from the CDD and InterPro searches of the CoaBC protein sequences were in agreement with the literature [209, 210]. As reported in literature, the PPCDC domain was predicted to be at the N-terminal end and the PPCS domain at the C-terminal end of the bifunctional protein. This was further supported by the respective signatures that were identified in the separate proteins of *M. iowae*, which match the regions of the two identified domain signatures. The signal peptide sequence that was predicted within most of the protein sequences suggested that these proteins are membrane-bound proteins on the outside of the membrane i.e. the extracellular environment. If this is the case, the enzyme

substrate, 4'-phosphopantothenate (**2.8**, Scheme 2.2), will have to be exported, following the reaction of PanK. This seems highly unlikely since this substrate is very polar (consisting of a phosphate group and a carboxylic acid moiety) and would, subsequently, require an allocated transporter to cross the membrane. Alternatively, 4'-phosphopantothenate might also be obtained from the host. However, this is merely speculation, as it would suggest that 4'-phosphopantothenate is readily available in the environment.

Although the motif regions predicted by MEME did not necessarily contain literature reported motif sequences of the respective protein domains, the literature reported motif sequences were still present in the identified protein sequences. Furthermore, the absence of Motif 4 in the CoaBC protein sequences of *M. mobile* and *M. testudinis* were both due to a large deletion. However, these deletions did not affect the presence of any literature reported motifs [210]. However, the absent motif in the PPCS protein sequence of *M. iowae* is the result of a completely different sequence. This observation is not unexpected since a single monofunctional protein is being compared to several bifunctional proteins. Therefore, it seems that the sequence similarity of the PPCS domain of CoaBC and the separate PPCS protein is not as high as that of the PPCDC domain and the separate PPCDC protein.

The CDD and InterPro predictions on the various PPAT protein sequences were in agreement with the literature, including the nucleotidyltransferase superfamily [148]. The shorter recognition region in the PPAT protein sequence of *M. mobile* is a possible reason for it being annotated as "putative". This shorter recognition region also corresponds to the MEME search results, in which Motif 3 and Motif 4 were predicted to be absent in the protein sequence of *M. mobile*. The inspection of the MSA of the PPAT protein sequences revealed that the C-terminal end of the *M. mobile* protein sequence displays low similarity to the other sequences. This low similarity did, however, not affect prominent residues associated with substrate binding [148]. Furthermore, since the MEME-predicted motifs in the PPAT protein sequences covered almost the entire length of the protein sequence, there were segments of these predicted motifs that agree with the literature reported motifs [148, 150], including prominent residues involved in substrate binding.

The results obtained from the InterPro and CDD searches of the DPCK protein sequences agree with the reported literature, concerning superfamily prediction [157]. The MEME-predicted Motif 1 within all the DPCK protein sequences revealed the presence of the Walker A motif [110], which corresponds with the literature [157]. Additionally, all of the motifs predicted for the HAD-DPCK protein sequences indicated that the DPCK domain of the protein is at the C-terminal end. This was also supported by the CDD and InterPro search results.

The CDD searches could only predict a general function for the HAD-like domain of the HAD-DPCK proteins. The specific hit was identified as Cof, which refers to the Cof protein of *Escherichia coli* (hydroxymethylpyrimidine pyrophosphatase) and other HAD family phosphatases. This Cof protein catalyses the hydrolysis reaction of 4-amino-2-methyl-5-hydroxymethylpyrimidine pyrophosphate (HMP-PP) to 4-amino-2-methyl-5-hydroxymethyl-pyrimidine phosphate (HMP-P), which are substrate intermediates in thiamine (vitamin B_1) metabolism [211]. In the same study by Lawhorn *et al.* [211], it was reported that this enzyme can also catalyse the hydrolysis of other substrates such as 4-amino-2-methoxy-5-hydroxymethylpyrimidine pyrophosphate and 4-amino-2-trifluoromethyl 5-hydroxymethyl-pyrimidine pyrophosphate (CF_3-HMP-PP) to give 4-amino-2-methoxy-5-hydroxymethyl-pyrimidine phosphate and 4-amino-2-trifluoromethyl 5-hydroxymethylpyrimidine phosphate, respectively. These hydrolysis reactions can, therefore, provide resistance to antibiotics such as bacimethrin and 4-amino-2-trifluoromethyl 5-hydroxymethyl-pyrimidine, by detoxifying their toxic forms, 2'-methoxythiamine pyrophosphate and CF_3-HMP-PP, respectively. Additionally, hydrolysis of secondary substrates, including pyridoxal-phosphate (the active form of vitamin B_6) and flavin mononucleotide (produced from vitamin B_2), as well as purines (GMP and IMP), have also been reported for Cof and other HAD phosphatases [212]. Hence, these types of phosphatases are promiscuous and seem to be associated with numerous vitamin and cofactor metabolic processes.

In addition to its potential antibiotic-resistant properties, a HAD phosphatase domain in the HAD-DPCK enzyme might be advantageous in mycoplasmas for a number of reasons. For instance, due to the promiscuity of the HAD phosphatase enzyme, the HAD-DPCK enzyme might have regulatory properties, such as regulation of the intracellular levels of various coenzymes. Furthermore, the fusion of these enzymes was most likely the result of degenerative evolution, in an attempt to optimise the genomic composition of these mycoplasmas.

Of the 62 investigated *Mycoplasma* species, there were eight species in which no CoA biosynthetic enzyme-encoding genes could be identified. Five of these *Mycoplasma* species form part of the hemotropic mycoplasma cluster (a division within the pneumoniae group) [26]. This cluster can also be seen in the 16S rRNA sequence-based phylogeny (Figure 3.1) by its distinctive clade within the pneumoniae group. The unique *Mycoplasma* species in this cluster (also referred to as hemoplasmas) are blood-borne pathogens that infect the erythrocyte [213]. This nutrient-rich environment might have allowed the hemoplasmas to undergo additional metabolic reduction (as a result of degenerative evolution) when compared to the other mycoplasmas. This is supported by a comparative genomics study of

the hemoplasmas with other *Mollicutes* species [214], which revealed the absence of all enzymes involved in the pentose-phosphate pathway, CoA metabolism, and the pyruvate dehydrogenase complex in hemoplasmas. Additionally, the latter two was also found to be absent in some members of the hominis group, which might explain the other three species with no identified CoA biosynthetic enzyme-encoding genes.

This absence of enzyme-encoding genes associated with the CoA biosynthetic pathway is supported by a recent study by Hutchison *et al.* [215] in which a synthetic *Mycoplasma* species with a minimal genome was developed. Herein, they demonstrate that the entire CoA biosynthetic pathway could be dispensed with altogether. Furthermore, a major portion of the listed unknowns in the description of this synthetic species resembles membrane proteins similar to transport proteins, albeit without clearly allocated functions. Therefore, it could be possible to source intact CoA directly from the environment.

The most abundant CoA biosynthetic enzyme-encoding gene that could be identified amongst the investigated mycoplasmas was the DPCK-encoding gene. Except for the hemoplasmas, *M. hominis*, *M. arthritidis*, *M. canadense* and *M. gallinaceum*, this gene was present in all of the mycoplasmas. This absent DPCK-encoding gene in *M. gallinaceum*, however, seems highly unlikely, especially since all of the other CoA biosynthetic pathway enzyme-encoding gene could be identified. There is currently only one published genome available of this organism [216]. However, the article has been retracted after it was discovered that the sequenced genome was in fact of *M. gallinaceum*, instead of *Mycoplasma meleagridis* as initially reported. Although this error was corrected on the NCBI database by changing the *Mycoplasma* species of the sequenced genome to *M. gallinaceum*, a revised manuscript has yet to be published. Subsequently, this error has raised questions about the reliability of the sequenced genome. Incomplete and error-filled genome assemblies are not uncommon and can lead to many annotation errors, including the determination of the number of genes present in a genome [217, 218]. This might possibly be the case with the *M. gallinaceum* genome, resulting in a false identification of an absent DPCK-encoding gene.

Eleven species were identified that only have an enzyme-encoding gene for the final step in the CoA biosynthetic pathway. Six of these were identified to be a DPCK-encoding gene, where two species are in the hominis group and the remaining in the pneumoniae group, according to the 16S rRNA sequence-based phylogeny (Figure 3.1). In the other five species, which are also in the hominis group, a HAD-DPCK-encoding gene was identified. Six of the hominis group species (with Ms02 as the exception) are located within the same clade in the 16S rRNA sequence-based phylogeny. However, the genome of Ms02 is

currently only available at the contiguous sequence level and the full genome still needs to be completed. Therefore, although the sequences were sufficiently complete to identify only the DPCK-encoding gene, the presence of the other CoA biosynthetic enzyme-encoding genes cannot be ruled out.

Although there were five species in which only the unknown HAD-DPCK protein-encoding gene could be identified, there was a sixth species that is located within the same clade (*M. conjunctivae*), in which a PPAT- and HAD-DPCK-encoding gene could be identified. However, there seems to be no apparent reason for this inconsistency, as all of the species with this identified HAD-DPCK protein infect similar hosts (livestock), have similar sources of energy (glucose) and have even been isolated from similar sources (conjunctivae, apart from *M. hyopneumoniae* and *M. dispar*).

The fact that only the final enzyme in the CoA biosynthetic pathway could be identified for the above eleven species, suggests that they acquire their CoA by scavenging dephospho-coenzyme A (DePCoA) from the host or the surrounding environment, possibly via an (as of yet) unidentified transport system. This observation has also been made in various *Rickettsia* and *Chlamydia* spp. [7, 219]. The obvious dependency of mycoplasmas on their host for many metabolites and intermediates, due to their restricted metabolic biosynthetic potential, highlights the necessity of a vast set of transporters. However, in mycoplasmas, many transporters are vaguely annotated on the basis of similarities to established transporter classes or membrane segments, yet little is known regarding their actual substrates [220]. Furthermore, the characterisation of a mitochondrial transporter of DePCoA in *Drosophila melanogaster* in a recent study by Vozza *et al.* [221] suggests that a prokaryotic version of such a transporter might exist.

A combination of PPAT- and DPCK-encoding genes could be identified in almost 40% of the investigated *Mycoplasma* species, which included all of the species within the spiroplasma group and none of the species within the pneumoniae group. Based on the 16S rRNA phylogeny, the species within the spiroplasma group seem to be very consistent regarding the number of CoA biosynthetic pathway enzyme-encoding genes present. This consistency can also be seen in the locations of these genes relative to each other, as well as their phylogenetic relationships. The majority of species within the hominis group with only these two enzyme-encoding genes are members of the *bovis* cluster [26]. Members of this cluster also seem to be very consistent with regards to the number of CoA biosynthetic pathway enzyme-encoding genes present. However, the other hominis group species are members of the *synoviae* cluster [26], which seem much more diverse in this regard.

This genomic CoA biosynthetic pathway variant, in which only the final two enzyme-encoding genes are present, is not as unlikely as it seems. Recently, it has become evident that cells, including some micro-organisms, might be able to utilise 4'-phosphopantetheine (PPantSH, **2.10**, Scheme 2.2) to fulfil their CoA needs [222]. This was supported in a study by Srinivasan *et al.* [223], which provided strong evidence that PPantSH is able to cross the membranes without the help of a transporter in *D. melanogaster* Schneider 2 cells. Whether this is the case in these *Mycoplasma* species and whether they have an allocated PPantSH transporter remains an open question. There is, however, no question regarding the availability of PPantSH. In addition to being an intermediate product of the *de novo* CoA biosynthetic pathway (required by the host), PPantSH is also produced via the degradation of CoA and the acyl carrier proteins involved in lipid biosynthesis [94]. Furthermore, *E. coli* was found to produce such a large amount of PPantSH that it was excreted, but also never re-imported [93]. Thus, whether PPantSH is obtained from the host or whether it is obtained from surrounding microbial communities, there is still an abundant supply of PPantSH and this suggests that these *Mycoplasma* species might indeed salvage PPantSH from the environment.

The PanK-encoding gene present in all eight species in which only the three salvage pathway enzyme-encoding genes (PanK, PPAT and DPCK) could be identified was revealed to code for the type III PanK enzyme. This also seems to be the case in the human pathogen *Treponema pallidum*, which causes syphilis [9]. Consequently, since the type III PanKs are unable to phosphorylate PantSH, this enzyme might be useless with respect to the biosynthesis of CoA via the salvage pathway. However, in light of the previously mentioned Pan Motif mutation observed in the identified PanKs of mycoplasmas, this enzyme might be able to accept PantSH as a substrate. Alternatively, these PanKs could have an entirely different role within these organisms since PanK enzymes are notoriously promiscuous concerning substrates. If this is indeed the case, PPantSH can then be salvaged as mentioned above in order to satisfy the CoA requirements.

Only nine of the investigated *Mycoplasma* species contained enzyme-encoding genes for the complete five-step CoA biosynthetic pathway. Interestingly, these species are widely distributed within the hominis- and pneumoniae phylogenetic groups, and not restricted to a single phylogenetic group/cluster according to their 16S rRNA sequences. There are also no obvious indications, with respect to the specific host and/or energy source, as to why only these species possess the entire pathway compared to the other investigated species. However, there does seem to be some consistency relating to the genomic locations of the

PanK-encoding gene and the CoaBC-encoding gene, with respect to each other, since these two genes have overlapping coding regions in all of the relevant species.

Additionally, *M. iowae* is the only species of the nine with separate PPCS- and PPCDC-encoding genes. This *Mycoplasma* species is well-known for the unique characteristic of phenotypic variation in the *Mycoplasma* surface components, in addition to a relative resistance to bile salts, heat, and many antimicrobials [224]. Subsequently, this deviation from the norm might be due to any (or a combination) of these differences. However, since InterPro predicted that most of the CoaBC enzymes are membrane-bound, it is most probably as a result of phenotypic variation.

The overall bootstrap support values of the phylogenies based on the respective proteins are reasonably low, especially the phylogenies based on the PPAT- and DPCK sequences in comparison to that of the 16S rRNA sequences. These low values could be due to the greater character variability associated with amino acid sequences (20 possibilities), compared to nucleotide sequences (4 possibilities). Consequently, the increased inconsistency within the informative characters of the amino acid sequences results in a lower reproducibility of a specific node in the bootstrap iterations, especially with larger data sets. In contrast, the smaller data sets still have enough redundancy within the informative characters to provide satisfactory values, as is the case in the PanK- and CoaBC sequence-based phylogenies. However, it is important to note that bootstrap analysis is not a test concerning the accuracy of the phylogeny; it merely gives an indication of the stability of the tree topology and also helps to assess whether the sequence data is sufficient enough to validate the topology [225]. Hence, although the topology of the DPCK sequence-based phylogeny (Figure 3.6) closely resembles that of the 16S rRNA sequence-based phylogeny (Figure 3.1), the bootstrap values still provide statistical support (or a lack thereof, in this case) and should, therefore, still be considered.

Accordingly, apart from the well-supported distinctive spiroplasma grouping, no significant conclusions could be made with respect to the phylogenetic relationships of the investigated *Mycoplasma* species, based on the PPAT- and DPCK protein sequences compared to that of the 16S rRNA sequences. Conversely, the phylogenies based on the PanK- and CaoBC sequences revealed that the functional relationship, of these respective enzymes, between the *Mycoplasma* species have no correlation with their corresponding phylogenetic relationship based on the 16S rRNA sequences. The MSAs, in combination with the MEME motif predictions, of the respective protein sequences gave an indication of the functional relationship between the species proteins, which occasionally corresponded with what is suggested by the respective protein sequence-based phylogeny.

Additionally, the ancestor node for *M. arginini* observed in the 16S rRNA sequence-based phylogeny compared to that observed in all of the protein sequence-based phylogenies is a noteworthy difference. In the phylogeny based on the 16S rRNA sequences, *M. arginini* shares an ancestor node with *Mycoplasma canadense*, wherein all the other phylogenies it shares one with *M. synoviae*, which is in an entirely different cluster. These nodes are also very well supported by high bootstrap values. Strangely, *M. arginini* and *M. synoviae* do not infect similar hosts (mammals vs. galliforms, respectively) nor do they utilise the same energy source (arginine vs. glucose, respectively).

The number of identified genes within a genome, which code for CoA biosynthetic pathway enzymes, seems to vary immensely amongst the investigated *Mycoplasma* species and although there are some similarities with the 16S rRNA sequence-based phylogeny, the number of identified genes does not correlate with the phylogenetic relationship of the majority of the investigated *Mycoplasma* species. The most prominent of the similarities were the consistency of identified CoA biosynthetic pathway enzyme-encoding genes in the *Mycoplasma* species within the spiroplasma group, as well as that of the hemoplasmas. Conversely, apart from the species within the *bovis* cluster (which demonstrated some consistency), the remaining investigated *Mycoplasma* species from the pneumoniae- and hominis groups mainly displayed a broad variation with respect to enzyme-encoding genes.

Therefore, there is unpredictability in the genomic variation concerning the number of CoA biosynthetic pathway enzyme-encoding genes within the genomes of the investigated *Mycoplasma* species. This unpredictability is further supported by the observation that the number of enzyme-encoding genes is independent of the host or energy source of the respective *Mycoplasma* species. The genomic variation of the CoA biosynthetic pathway, therefore, seems to be entirely dependent on the needs of the individual species.

Nevertheless, there is still potential for the CoA biosynthetic pathway as a drug/vaccine target in some mycoplasmas, as supported by Balish *et al*. [64]. Be that as it may, a comprehensive investigation into the CoA biosynthetic pathway of the targeted mycoplasma will have to be performed in order to identify the genetic variation present within the genome of the specific species. Furthermore, eight species, including the hemoplasmas, were demonstrated to possess none of the enzyme-encoding genes, which suggested the potential dispensableness of this pathway. However, this only amounts to 13% of the total investigated *Mycoplasma* species, of which 8% are the hemoplasmas. In the remaining species, the most consistently present enzyme-encoding gene was that of DPCK. Therefore, DPCK seems to be the most important enzyme involved in the biosynthesis of CoA in the

investigated mycoplasmas, and should, consequently, be the starting point in the investigation into anti-mycoplasmal agents.

Chapter 4 – Cloning, expression and isolation of the *Mycoplasma* sp. Ms02 dephospho-coenzyme A kinase

4.1 Introduction

In light of the results presented in the previous chapter, the coenzyme A (CoA) biosynthetic pathway might be a possible drug/vaccine target within ostrich-infecting mycoplasmas. More specifically, the dephospho-coenzyme A kinase (DPCK), being the most abundantly present of all the pathway enzymes amongst mycoplasmas, could be a promising start. The genome of *Mycoplasma* sp. Ms02 (Ms02), which is available and annotated (albeit at contiguous sequence level) [57], was sufficient enough to identify its DPCK-encoding gene (*msDPCK*). Consequently, the aim of this study was to express and isolate the Ms02 DPCK enzyme for the purpose of future characterisation and the subsequent comparison to characterised DPCK enzymes of other organisms. To this end, *msDPCK* was amplified and subsequently cloned into a suitable expression vector. Next, the expression and isolation of this DPCK enzyme were attempted and finally, preliminary tests performed to determine enzymatic activity.

4.2 Background on experimental approaches

This section will provide additional background information on some of the experimental procedures utilised in this study to achieve the above-mentioned objectives. Furthermore, due to solubility issues with the Ms02 DPCK enzyme, the expression and isolation of the protein proved to be problematic. This prompted an investigation into alternative methods resulting in various attempts at achieving solubility. The different approaches that were used for these solubility attempts are, therefore, also discussed.

4.2.1 Experimental procedures

4.2.1.1 Site-directed mutagenesis

Members of the class Mollicutes utilise a different genetic code in comparison to other organisms. The universal usage of the UGA codon is one of three stop codons; whereas in Mollicutes the UGA codon encodes for a tryptophan [22, 33]. This can, therefore, be problematic for recombinant expression of mycoplasma genes that contain such codons in *Escherichia coli*, since it will lead to the premature termination of the expressed protein [226]. One of the ways to overcome this problem is to use an *E. coli* strain that provides tRNA suppressors of the UGA codon [227]. However, this method does not work very well with genes containing more than one UGA codon. Another alternative is to use an

expression system that utilises the same genetic code, such as *Spiroplasma citri*, but these organisms are difficult to cultivate [228]. Last but not least, is to perform site-directed mutagenesis (SDM) [229], which was the preferred method for the purpose of this study.

SDM can be used to change the TGA codons (DNA equivalent of the UGA codon) to TGG codons (universal tryptophan codon) within the relevant mycoplasma gene. This mutation is created by performing a polymerase chain reaction (PCR) with specifically designed primers that contain the mutated sites. The targeted gene, which is already cloned into a vector, is used as the template DNA for this PCR (Figure 4.1). This allows the target gene and vector to be amplified and, subsequently, produces a vector containing the mutated target gene. The original template vector is then removed by digestion with DpnI endonuclease, which digests methylated and hemimethylated DNA. Consequently, the mutated vector plasmid is transformed into the desired *E. coli* strain, followed by confirmation of the mutated sites via sequencing.

The sequence of the vector should, however, also be confirmed as errors during the PCR amplification may occur, leading to unwanted mutations. In order to avoid continuing with a vector plasmid containing possibly unwanted mutations, the correctly mutated target gene insert is sub-cloned into a new vector, followed by the final confirmation of the mutated sites via sequencing.

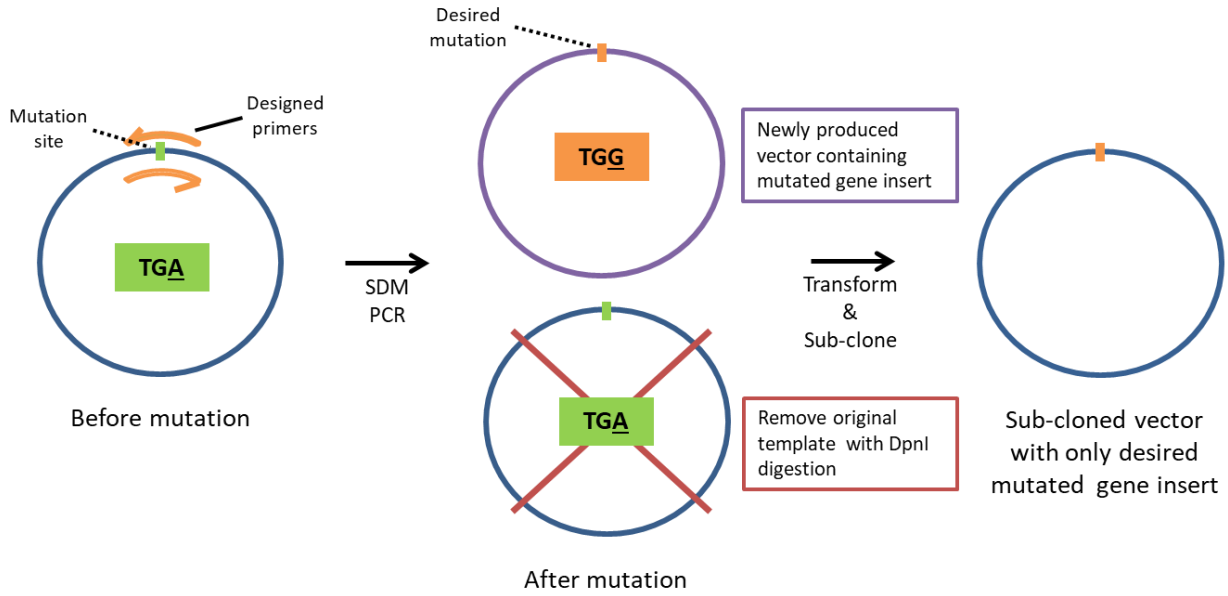


Figure 4.1 A simplified graphic illustration of the site-directed mutagenesis procedure.

4.2.1.2 Primer design

There are general criteria for designing the ideal primer pair [230]. Firstly, the length of the oligonucleotides should be in the range of 15-24 nucleotides. If they are too short, they can lack specificity; if they are too long their annealing efficiency might be reduced. Additionally, the two primers should not differ by more than five nucleotides. Similarly, the estimated annealing temperatures (T_m) of the two primers must preferably be within the range of 55-65˚C and not differ with more than 5˚C of each other. If possible, the G+C content of the primers should be in the range of 55-60%, in which long runs (repeats) of these nucleotides should be avoided. Furthermore, primer binding may be promoted by the presence of a G or C nucleotide at the 3'-end of the primer (known as the G/C-clamp). Finally, primer self-complementarity (i.e. hairpin loops or primer dimers) should be avoided where possible.

The above criteria are, however, merely guidelines and may differ depending on the organism in which they are being designed, or the circumstances for which they are being designed. For example, in mycoplasmas, it can be difficult to fulfil all of these criteria, specifically the G+C content, as mycoplasma genomes have a naturally low G+C content. Moreover, primers were designed for the amplification of *msDPCK* for cloning into various expression vectors, as well as for SDM. The designed primers for cloning purposes required the addition of the relevant restriction endonuclease (RE) site at their 5'-end, along with a leader sequence of 4-5 random nucleotides that assists with efficient RE digestion. Conversely, the primers for SDM are designed with the site of mutagenesis positioned more or less in the middle, with 18-24 nucleotides flanking both sides of the site. Therefore, these primers tend to be very long. They also have a degree of primer self-complementarity, especially primer dimers, as these primers are designed on complement strands of the same position.

4.2.1.3 Enzyme activity evaluation

In order to determine if the isolated Ms02 DPCK enzyme was active, a high-performance liquid chromatography (HPLC) based method was used, as reported by Goosen *et al*. [231]. This method allows the simultaneous quantification of CoA and its thiol-carrying precursors with pmol sensitivity. This is achieved by utilising a thiol-derivatisation agent that produces a fluorophore upon labelling, followed by separation thereof via HPLC and quantification with fluorescence detection. Therefore, the simultaneous detection of dephospho-coenzyme A (DePCoA, substrate of DPCK) and CoA (product of DPCK) can be used to determine if the isolated enzyme demonstrated enzymatic activity, i.e. if it was able to convert DePCoA to CoA within an *in vitro* enzyme reaction mixture.

4.2.2 Experimental approaches for attempts at solubility

4.2.2.1 Shifting the terminal position of the 6xHis-tag

Initially, soluble expression attempts were performed with a 6xHis fusion tag at the N-terminus of the target protein to allow purification by means of immobilised metal affinity chromatography (IMAC). However, it has been reported that a His-tag at the N-terminus of a protein might be detrimental to the solubility, folding or oligomerisation properties of the particular protein [232]. Moreover, His-tags at the N-terminus and C-terminus of a protein can have different influences on the protein structure and properties [233]. Therefore, the His-tag was shifted from the N-terminal end to the C-terminal end, with the notion that it will improve the solubility of the target protein.

4.2.2.2 Modified MBP-6xHis-tag

It is well-known that certain affinity tags are able to increase the solubility of some of the proteins they are attached to [234, 235]. The most notable of these are the maltose-binding protein (MBP) and the glutathione-*S*-transferase (GST) affinity tags. However, the latter has since been reported to be quite poor with regards to enhancing the solubility of its attached protein partner [236, 237]. On the other hand, several studies have reported soluble protein production subsequent to its fusion with an N-terminal MBP, where the unfused protein equivalent demonstrated an insoluble product [237–240]. This is also the only natural affinity tag that has been thoroughly validated as a solubility enhancer [241]. For these reasons, the fusion of the target protein with an MBP-tag was selected as an approach to enhance its solubility.

To this end, the vector selected for in this study was a modified MBP-6xHis-tag vector (pMALcHT, described by Muench *et al*. [242]), in which the MBP-tag and 6xHis-tag are separated by a tobacco etch virus (TEV) protease cleavage site (Figure 4.2). Thus, the co-expression with a plasmid encoding TEV protease (in this study pRK586, described by Kapust *et al*. [243]) will allow the *in vivo* cleaving of the MBP-domain, resulting in a 6xHis-tag protein, which can subsequently be purified by means of IMAC. The plasmid used for sub-cloning of *msDPCK* in this study was the pSPr022 plasmid (Figure 4.2), which was acquired from Prof. Sean T. Prigge (Johns Hopkins University) and has been described in Du *et al.* [244]. This is essentially the pMALcHT plasmid with a *Plasmodium falciparum* acyl carrier protein insert (inserted using RE digestion with EcoRI and SalI). Therefore, the *P. falciparum* acyl carrier protein (ACP) gene insert needed to be removed and the mutated *msDPCK* gene sub-cloned into the subsequent pMALcHT plasmid using the same RE sites.

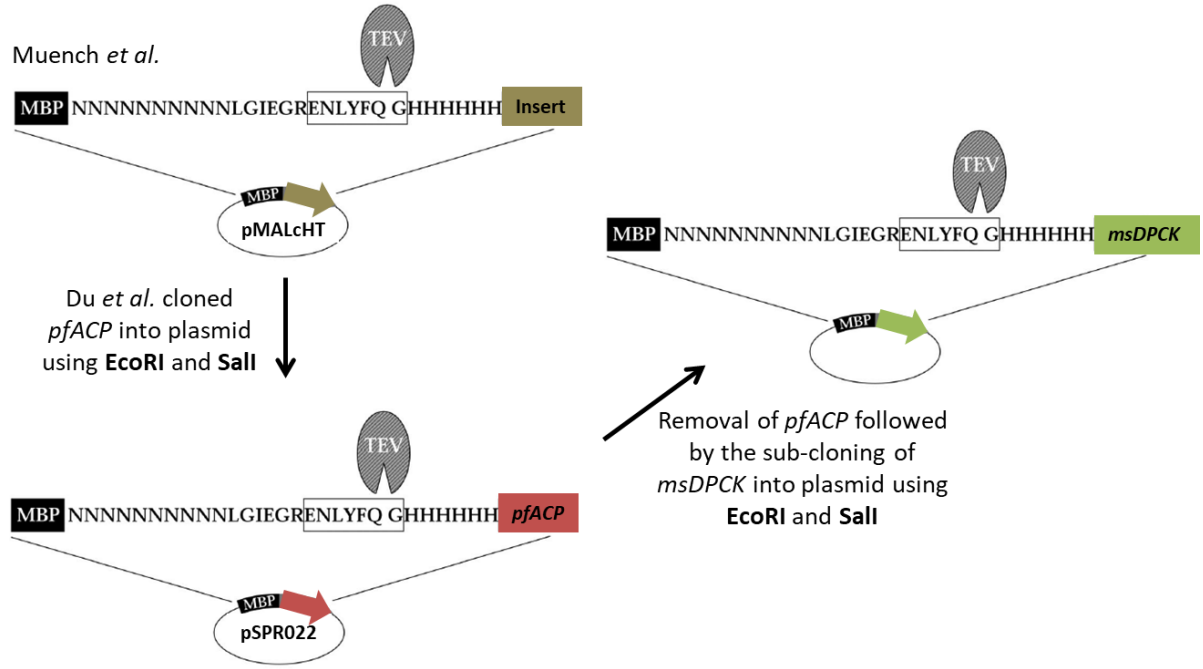


Figure 4.2 A schematic representation of creating of the modified MBP-6xHis-tag plasmid construct used in this study. Adapted from Muench *et al.* [242].

4.2.2.3 Molecular chaperones

The co-expression of molecular chaperones has been reported to improve the solubility of expressed recombinant proteins in numerous studies [245–249]. Since certain molecular chaperones are involved in protein folding [248], these chaperones could work cooperatively in this process and prevent the formation of inclusion bodies. Therefore, co-expression with these molecular chaperones might improve the soluble recovery of the expressed target protein. The pG-KJE8 plasmid (supplied by TaKaRa) that was used in this study, expresses the DnaK, DnaJ, GrpE, GroES, and GroEL molecular chaperones (Figure 4.3).

4.2.2.4 Sarkosyl detergent treatment

Sarkosyl (also known as N-laurylsarcosine) is a mild detergent with a negatively charged carboxylate group similar to sodium dodecyl sulphate, but it does not have the same denaturing outcome. It has been used in protein purification for over 40 years after it was first shown to solubilise proteins from bacterial cytosolic membranes, as well as folded recombinant proteins from inclusion bodies [250, 251]. More recently, a simple and effective method of using sarkosyl to solubilise troublesome recombinantly expressed proteins has been described [252]. This protocol reported that the addition of 1-2% sarkosyl to the lysis buffer during the lysis step, subsequent to expression, will almost certainly solubilise the protein in question; however, if the 2% sarkosyl is insufficient, soaking the post-lysate pellet in 10% sarkosyl will surely solubilise the proteins from inclusion bodies (>95%).

Furthermore, sarkosyl-solubilised 6xHis-tagged fusion proteins can reportedly be directly purified on Ni^{2+} resin columns, and the protein purified by this method can be used for mass spectrum assays, structure analysis, and biological assays. Therefore, this method was used as an attempt to produce soluble fractions of the target protein.

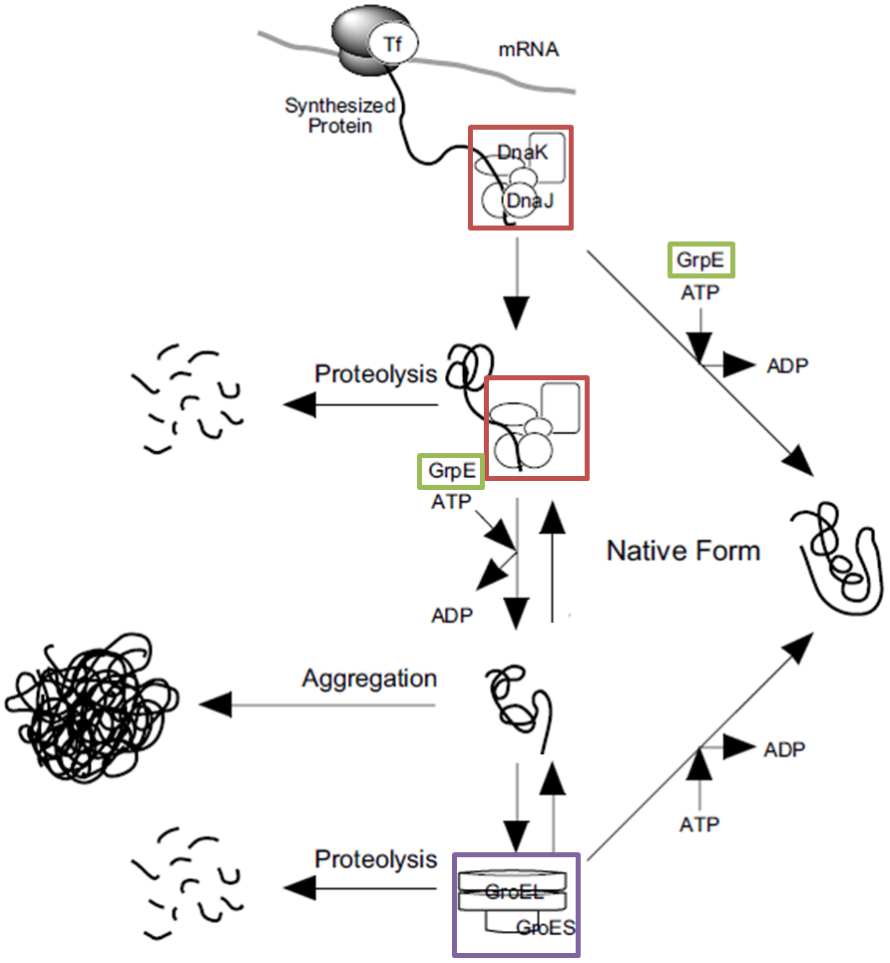


Figure 4.3 A schematic representation of the proposed model for chaperone-assisted protein folding in *E. coli*. DnaJ binds to the ribosome emerging polypeptides that target newly produced proteins for recognition by DnaK. The DnaJ-DnaK system (red box) protects stretches of hydrophobic residues in the polypeptide chain, which prevents unwanted interactions with cellular components and other intermediates. Binding of the GrpE nucleotide exchange factor (green box) to DnaK catalyses the release of the bound protein substrate, this may result in the substrate: folding into its native form; being transferred to the GroEL-GroES system (purple box); misfolding and aggregating; or undergoing additional cycles of DnaJ-DnaK-GrpE binding and release. The GroEL-GroES system is believed to facilitate the favourable isomerisation of proteins that are already in an intermediate conformation. If the protein released from the GroEL-GroES complex is unable to fold into its native form, it might aggregate of undergo another interaction cycle with either chaperone system. Both chaperone systems might also present their substrates to the cellular protease machinery. The binding and release of protein substrates with both DnaK and GroEL are ATP-dependent reactions that are coordinated by interactions with the DnaJ-GrpE or GroES cofactors, respectively. Adapted from Thomas *et al.* [245].

4.3 Materials and Methods

4.3.1 Design and treatment of primers

4.3.1.1 Primer design

All primers used in this study were designed using the computer program Primer Designer v1.01 (Scientific & Educational Software) and were synthesised and purified by Integrated DNA Technologies, USA. The *msDPCK* gene (579 bp), obtained from the previously annotated genome [57], was used as template. The appropriate RE sites, with accompanying leader sequences, were included in each of the primers used for cloning or sub-cloning (Table 4.1).

4.3.1.2 Treatment of primers

Following initial primer arrival, the primers were dissolved in 1x TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0) to provide a primer stock (concentration of 100 µM), which was stored overnight at 4˚C to hydrate. The following day the primers were incubated at 60˚C for 15 min, cooled to 4˚C, and finally stored at -20˚C. Subsequently, this primer stock was used to make aliquots of working stock (10 µM).

Table 4.1 The designed primers used for cloning or sub-cloning

Primer set	Name	Optimal T_a[a]	RE site	Primer sequence[b]
1	DPCK02_F*	55˚C	NdeI	5'-AGGTCATATGATGATTGCTATTACAGGATTACTAAAAG-3'
	DPCK02_R**		XhoI	5'-GGATCTCGAGTTATTTTAAATTAGGGCAAATATATTTCG-3'
2	DPCK_C-term_F*	50˚C	NcoI	5'-AGGTCCATGGATGATTGCTATTACAGGATTACTAAAAG-3'
	DPCK_C-term_R**		XhoI	5'-GGATCTCGAGTTTTAAATTAGGGCAAATATATTTCG-3'
3	DPCK_MBPf*	50˚C	EcoRI	5'-AGGTAGAATTCATGATTGCTATTACAGGATTACTAAAAG-3'
	DPCK_MBPr**		SalI	5'-AGGTAGTCGACTTATTTTAAATTAGGGCAAATATATTTCG-3'

*Forward direction primer
**Reverse direction primer
[a]Optimal annealing temperature determined by temperature gradient PCR
[b]Restriction endonuclease site is underlined, with leader sequence at its 5'-end

4.3.2 Amplification, cloning and site-directed mutagenesis of *msDPCK*

4.3.2.1 Amplification and isolation of msDPCK

Using the designed Primer set 1 (Table 4.1) the *msDPCK* gene was amplified by PCR from Ms02 genomic DNA (previously isolated by Steenmans [57]). The following reaction conditions were used: 5 µl of 5x KAPA HiFi buffer (Fidelity), 0.75 µl of KAPA deoxynucleotide triphosphates (dNTPs, 10 mM), 0.75 µl of each primer (Primer set 1, 10 µM), 2 µl of template DNA (110.2 ng/µl genomic DNA diluted to 1:50 in Milli-Q® water), 0.5 µl of KAPA HiFi DNA Polymerase (1 U/µl, KAPA Biosystems, RSA) and Milli-Q® water to

a final volume of 25 µl. The amplification parameters set on the Veriti™ 96 Well Thermal Cycler (Applied Biosystems) consisted of an initial step of 3 min at 95˚C, followed by 30 cycles of 20 sec at 98˚C, 15 sec at optimal annealing temperature (T_a) and 90 sec at 72˚C with a final elongation step of 5 min at 72˚C, followed by cooling to 15˚C. The optimal T_a (Table 4.1) was determined by performing a temperature gradient PCR in increments of 5˚C (each temperature performed in duplicate).

The PCR amplifications were analysed on 2% (m/v) agarose gels prepared with SeaKem®LE Agarose (Lonza, Switzerland) dissolved in 1x TAE buffer (40 mM Tris-HCl, 20 mM acetic acid and 1 mM EDTA, pH 8). The gels contained 0.5 µg/ml ethidium bromide (Sigma-Aldrich, USA) for visualisation of PCR products under ultraviolet light. Samples were mixed (5:1) with loading dye (0.1% (v/v) bromophenol blue, 50% (v/v) glycerol, 100 mM Tris-HCl and 50 mM EDTA, pH 8) and electrophoresis was performed in 1x TAE buffer at a constant voltage of 100 V. The 100 bp DNA ladder (Promega, USA) was used as marker (5 µl loaded).

In order to obtain a larger yield for purification and subsequent RE digestion purposes, eight PCR amplifications (using the same parameters) were performed and subsequently pooled together to produce a final volume of 200 µl. The pooled product was purified using the Wizard®SV Gel and PCR Clean-Up System (Promega) according to the manufacturer's instructions. The concentration and purity of the purified product were determined with the use of a NanoDrop® ND-1000 v3.5.1 spectrophotometer (Novell®, USA). The purified product was also analysed on an agarose gel, as previously described, to confirm that the desired PCR product was in fact purified.

4.3.2.2 Restriction enzyme digestion of msDPCK *PCR product and pET28a(+) producing N-terminal 6xHis-tag*

Both the *msDPCK* PCR product (amplified using Primer set 1, Table 4.1 and Table 4.2) and pET28a(+) plasmid vector containing N-terminal 6xHis-tag were subjected to double digestion using the NdeI and XhoI REs. Each reaction mixture contained: 2 µl of 10x CutSmart® buffer (New England BioLabs®, USA), 1 µl of each RE, the required volume of DNA (plasmid or insert) to give a final concentration of 50 ng/µl and Milli-Q® water to a final volume of 20 µl. The reaction mixtures were incubated at 37˚C for 25 min and then placed on ice before purification. Enzymes were removed by purification of digested samples using the DNA Clean & Concentrator™-5 Kit (Zymo Research, USA) according to the manufacturer's instructions. Subsequently, the digested plasmid was dephosphorylated using FastAP™ Thermosensitive Alkaline Phosphatase (Thermo Fisher Scientific, RSA) according to the manufacturer's instructions; with one deviation, i.e. no 10x Reaction buffer was added in the reaction mixture and instead was replaced with Milli-Q® water.

Table 4.2 The designed primer sets with their resulting plasmids including the parent vector and inserts that were used to construct them

Primer set	Plasmid	Fusion tag	Parents	Source of parent vector	Method of construction
1	pET28a(+)-*msDPCK*n	6xHis (N-terminal)	pET28a(+), *msDPCK* PCR product[a]	Novagen, RSA	Restriction digest and ligation
2	pET28a(+)-*msDPCK*c	6xHis (C-terminal)	pET28a(+), *msDPCK* PCR product[b]	Novagen, RSA	Restriction digest and ligation
3	pMALcHT-*msDPCK*	Modified MBP-6xHis	pSPr022, *msDPCK* PCR product[b]	Sean T. Prigge Described in Du *et al.* [244]	Restriction digest and ligation

[a]PCR product from Ms02 genomic DNA
[b]PCR product from SDM_pET28a(+)-*msDPCK*n plasmid

4.3.2.3 Cloning of msDPCK *into pET28a(+) containing an N-terminal 6xHis-tag*

The *msDPCK* insert was ligated into the digested and phosphorylated pET28a(+) plasmid by employing an insert to vector ratio of 1:1 and 3:1. The ligation reactions contained: 1 µl 10x Ligase buffer, 1 µl T4 DNA ligase (3 U/µl, Promega), the required volume of vector DNA to give a final concentration of 10 ng/µl, the calculated volume of insert DNA needed for the applicable ratio (Equation 4.1) and Milli-Q® water to a final volume of 10 µl. The ligation reactions were incubated overnight (O/N) at 4˚C.

Equation 4.1 Calculation for the amount of insert DNA (ng) needed for a particular ratio

$$\frac{\text{ng of vector} \times \text{kb size of insert}}{\text{kb size of vector}} \times \text{insert: vector molar ratio} = \text{ng of insert}$$

The ligated plasmids were transformed into *E. coli* JM109 competent cells (Promega). Before the transformation, these competent cells (stored at -80˚C) were placed on ice until properly thawed. The transformation was performed as follows: 50 µl of competent cells were added into a chilled 15 ml Falcon® tube, followed by the addition of 5 µl of the relevant ligation product. The contents of the tube were mixed by gently flicking it and then placed on ice for 20 min. Next, the cells were heat-shocked by submerging the tube (a quarter of the way) in a water bath (42˚C) for 45 sec, followed by immediately placing it back on the ice for 2 min. Subsequently, 950 µl of Luria Bertani (LB) medium (10 g/l Bacto-Tryptone, 5 g/l yeast extract, 10 g/l NaCl, pH 7) was added to the tube and incubated at 37˚C for 90 min on an orbital shaker (150 rpm).

The incubated cells were then transferred from the 15 ml Falcon® tube to a 1.5 ml microcentrifuge tube and the cells pelleted by centrifugation at 8 000 x g for 4 min. The transformed cells were then concentrated by removing 850 µl of the supernatant, followed by resuspending the pellet in the remaining supernatant. Finally, the total volume of this

transformed cell suspension (±150 µl) was plated on an LB agar plate (15 g/l), supplemented kanamycin (30 mg/l, Sigma-Aldrich), and incubated O/N (16-18 h) at 37˚C.

4.3.2.4 Colony PCR

Colonies that formed on the LB agar plates were analysed by performing colony PCR, in order to confirm the presence of the *msDPCK* insert. The following reaction conditions were used: 1 µl of 10x Reaction buffer, 0.6 µl of $MgCl_2$ (25 mM), 0.4 µl of dNTPs (5 mM, KAPA Biosystems, RSA), 1 µl of each primer (10 µM) of Primer set 1 (Table 4.1) or the T7 primer set (Table 4.3, Inqaba Biotec™, RSA), 0.1 µl of Super-Therm Taq DNA polymerase (5 U/µl, JMR Holdings, USA) and Milli-Q® water to a final volume of 10 µl. The colonies chosen for confirmation were scraped with a sterile toothpick and mixed into the 10 µl reaction mixture before PCR amplification. The amplification parameters set on the Veriti™ 96 Well Thermal Cycler consisted of an initial step of 5 min at 95˚C, followed by 25 cycles of 30 sec at 94˚C, 15 sec at 55˚C and 30 sec at 72˚C with a final elongation step of 7 min at 72˚C, followed by cooling to 15˚C. The PCR products were analysed by performing agarose gel (2%) electrophoresis as described previously.

Table 4.3 The T7 primer set used in the colony PCR and sequencing PCR for pET28a(+) plasmids

Name	Direction	Primer sequence
T7 promoter	Forward	5'-TAATACGACTCACTATAGGG-3'
T7 terminator	Reverse	5'- GCTAGTTATTGCTCAGCGG-3'

Colonies that tested positive for the insert were selected to make O/N cultures. This was done by inoculating 5 ml LB medium, supplemented with kanamycin (30 mg/l), with the selected colony (using a sterile toothpick) in a 15 ml Falcon® tube. This culture was incubated O/N at 37˚C on an orbital shaker (150 rpm). Subsequently, freezer stocks were made of the respective transformed *E. coli* JM109 cell cultures by adding 500 µl of O/N culture to 500 µl of 80% glycerol (1:1 ratio) in a 1.5 ml microcentrifuge tube and stored at -80˚C. The plasmids contained within these cultures are now referred to as pET28a(+)-*msDPCK*n.

4.3.2.5 Site-directed mutagenesis of msDPCK

There were two TGA codons within the 579 bp nucleotide sequence of *msDPCK* that required SDM (to produce TGG codons). The nucleotide positions of the TGA codons were 525 and 531. These sites were close enough to each other to allow simultaneous mutagenesis of both sites with one primer set (Table 4.4).

Table 4.4 The primer set designed for SDM of two sites in *msDPCK*

Name	Direction	Primer sequence[a]
SDM_DPCK_f1	Forward	5'-CTCATCAGGTAATCTGAGATGGAAATGGTATTTTAGAAAAATTAAAACG-3'
SDM_DPCK_r1	Reverse	5'-CGTTTTAATTTTTCTAAAATACCATTTCCATCTCAGATTACCTGATGAG-3'

[a]The mutated sites are underlined

The pET28a(+)-*msDPCK*n plasmid (Table 4.2) was isolated from O/N cultures, which was made using the *E. coli* JM109 freezer stocks (20 µl of freezer stock added instead of using a sterile toothpick). Plasmid isolation was performed using the Invisorb® Spin Plasmid Mini Two Kit (Invitek, USA) according to the manufacturer's instructions with one deviation, i.e. the plasmid DNA was eluted with 30 µl of pre-heated Milli-Q® water (65˚C). The concentration and purity of the isolated plasmid were determined with the use of a NanoDrop® ND-1000 v3.5.1 spectrophotometer (Novell®, USA).

The following reaction conditions were used for the SDM PCR: 10 µl of 5x KAPA HiFi buffer (Fidelity), 1.5 µl of KAPA dNTPs (10 mM), 1.5 µl of each primer (Table 4.4, 10 µM), 1 µl of template DNA (10 ng/µl isolated pET28a(+)-*msDPCK*n plasmid DNA), 0.5 µl of $MgCl_2$ (25 mM), 1 µl of KAPA HiFi DNA Polymerase (1 U/µl, KAPA Biosystems) and Milli-Q® water to a final volume of 50 µl. The amplification parameters set on the Veriti™ 96 Well Thermal Cycler consisted of an initial step of 2 min at 95˚C, followed by 20 cycles of 20 sec at 98˚C, 15 sec at 64˚C and 2 min 30 sec at 68˚C with a final elongation step of 4 min at 68˚C, followed by cooling to 15˚C.

The SDM PCR products were treated with DpnI for the digestion of the template plasmid. The reaction mixture contained: 5 µl of 10x Reaction buffer, 10 µl DNA (SDM PCR products), 0.1 µl Fermentas DpnI (10 U/µl, Thermo Fisher Scientific) and Milli-Q® water to a final volume of 50 µl. The mixture was incubated at 37˚C for 2 h, followed by inactivation of the enzyme at 80˚C for 20 min. Subsequently, the DpnI digested products were purified using the DNA Clean & Concentrator™-5 Kit (Zymo Research) according to the manufacturer's instructions and then transformed into *E. coli* JM109 competent cells (all 10 µl of the purified product was added) as previously described. The colonies were analysed by colony PCR, using Primer set 1 (Table 4.1), as previously described. Colonies that tested positive for the insert were selected to make O/N culture, followed by plasmid isolation as previously described.

The post-SDM isolated plasmids were sequenced to confirm the mutated sites within the *msDPCK* insert (DNA sequencing described in the following section). The plasmids with the complete *msDPCK* insert sequence and correctly mutated sites were subjected to RE double digestion (with NdeI and XhoI), followed by ligation of this insert into a new

pET28a(+) vector as described previously. These ligation products were transformed into *E. coli* JM109 competent cells and the subsequent colonies were analysed by colony PCR as previously described. Consequently, freezer stocks of the O/N cultures from the positively confirmed colonies (containing the SDM_pET28a(+)-*msDPCK*n plasmid) were made as described previously.

4.3.2.6 DNA sequencing

Before sequencing could be performed the relevant plasmid had to be isolated from O/N cultures grown from colonies as described previously; or alternatively grown form relevant freezer stocks. The *msDPCK* insert of the relevant isolated plasmid was sequenced by means of sequencing PCR (ABI BigDye® Terminator Cycle Sequencing Kit, Applied Biosystems). The following reaction conditions were used: 4 µl of BigDye, 3 µl of forward or reverse primer (3.3 µM) and 3 µl of isolated plasmid DNA (30 ng/µl). The forward or reverse primer of the T7 primer set (Table 4.3) was used. The amplification parameters set on the Veriti™ 96 Well Thermal Cycler consisted of 35 cycles of 10 sec at 96˚C, 30 sec at 52˚C and 4 min at 60˚C, with a final elongation step of 10 min at 60˚C and cooling to 15˚C. The PCR products were analysed by the sequencing facility of the Central Analytical Facility of the University of Stellenbosch, RSA.

4.3.3 Amplification and sub-cloning of SDM_*msDPCK*

*4.3.3.1 Amplification and isolation of SDM_*msDPCK

The amplification and isolation were performed as described in Section 4.3.2.1. However, in the PCR amplification reaction mixture, the relevant designed Primer set (Table 4.1) was used and the 2 µl Ms02 genomic template DNA was substituted for SDM_pET28a(+)-*msDPCK*n plasmid template DNA (137.8 ng/µl plasmid DNA diluted to 1:1000 in Milli-Q® water).

4.3.3.2 Restriction enzyme digestion

C-terminal 6xHis-tag – Both the SDM_*msDPCK* PCR product (amplified using Primer set 2, Table 4.1 and Table 4.2) and pET28a(+) plasmid vector containing a C-terminal 6xHis-tag were subjected to double digestion using the NcoI and XhoI REs. The reaction mixtures, purification thereof and dephosphorylation of the digested plasmid was performed as described in Section 4.3.2.2.

Modified MBP-6xHis-tag – The pSPr022 plasmid was previously transformed into *E. coli* Mach1 cells (Thermo Fisher Scientific) and cultures were stored at -80˚C. A freezer stock was used to prepare an O/N culture as described previously. However, the 5 ml LB medium

was supplemented with ampicillin (100 mg/l, Sigma-Aldrich). Plasmid isolation was subsequently performed as described previously.

Both the SDM_*msDPCK* PCR product (amplified using Primer set 3, Table 4.1 and Table 4.2) and pSPr022 plasmid vector were subjected to double digestion using the EcoRI and SalI REs. The reaction mixtures contained: 2 µl of 10x Fermentas FastDigest™ buffer (Thermo Fisher Scientific), 1 µl of each RE, the required volume of DNA (plasmid or insert) to give a final concentration of 50 ng/µl and Milli-Q® water to a final volume of 20 µl. The reaction mixtures were incubated at 37˚C for 15 min and then placed on ice before purification. The purification of the digested samples and dephosphorylation of the digested plasmid was performed as described in Section 4.3.2.2. The *P. falciparum* ACP gene insert was removed during purification of the plasmid producing the original linear pMALcHT plasmid, which was analysed via 1% (m/v) agarose gel analysis. The preparation and electrophoresis of the agarose gel were performed as described previously. The 1 kb DNA ladder (GeneDireX Inc., Taiwan) was used (5 µl) as a marker.

4.3.3.3 Ligation and transformation into E. coli *JM109 cells*

The ligation and transformation was performed as described in Section 4.3.2.3. However, the transformed cell suspension was plated on an LB agar plate supplemented with the relevant antibiotic, i.e. kanamycin (30 mg/l) for the pET28a(+)-*msDPCK*c (C-terminal 6xHis-tag) plasmid and ampicillin (100 mg/l) for the pMALcHT-*msDPCK* (Modified MBP-6xHis-tag) plasmid.

4.3.3.4 Colony PCR

A colony PCR was performed according to Section 4.3.2.4 in which the primers in the reaction mixture was substituted by the relevant primer set (Table 4.1), i.e. Primer set 2 for the pET28a(+)-*msDPCK*c plasmid and Primer set 3 for the pMALcHT-*msDPCK* plasmid.

4.3.3.5 DNA sequencing

The DNA sequencing of the gene inserts were performed as previously described. Similar to the pET28a(+)-*msDPCK*n plasmid, the primers of the T7 primer set (Table 4.3) was used in the reaction mixture of the pET28a(+)-*msDPCK*c plasmid. However, in the reaction mixture for the sequencing PCR of the pMALcHT-*msDPCK* plasmid, the primers of Primer set 3 (Table 4.1) was used.

4.3.4 Transformation of plasmids prior to expression

*4.3.4.1 Transformation of SDM_*msDPCK *containing plasmids into an* E. coli *expression strain*

For the purpose of protein expression, each of the plasmid constructs were respectively transformed into Invitrogen™ *E. coli* BL21 Star™ (DE3) cells (Thermo Fisher Scientific). To this end, the relevant plasmid was isolated from O/N cultures (prepared with the corresponding *E. coli* JM109 freezer stock) as described previously. However, before transformation of the plasmids into these cells they first had to be made competent. The chemical competent cells were produced as follows:

The Invitrogen™ *E. coli* BL21 Star™ (DE3) cells (stored at the -80˚C) were allowed to thaw on ice. By using a sterile inoculation loop, the thawed cells were streaked out onto an LB agar plate (no antibiotic) and incubated O/N at 37˚C. Subsequently, a single colony was inoculated in a 15 ml Falcon® tube containing 5 ml LB medium (no antibiotic) and incubated at 37˚C for 2 h. The contents of the Falcon® tube was then transferred to a 15 ml centrifuge tube, followed by centrifugation at 4 500 x g for 5 min. The supernatant was discarded and the pellet gently resuspended in 3 ml cold $CaCl_2$ (under sterile conditions), after which it was placed on ice for 30 min. This was followed by centrifugation at 4 500 x g for 5 min (4˚C). The supernatant was discarded and the pellet gently resuspended in 1 ml cold $CaCl_2$. These chemical competent cells were stored at 4˚C (not longer than 2 days) until used.

The appropriate plasmids were transformed into 80 µl of the chemical competent Invitrogen™ *E. coli* BL21 Star™ (DE3) cells as described previously. The subsequent colonies were analysed by colony PCR as before. Selected colonies were used to prepare O/N cultures, which was used for plasmid isolation and subsequent freezer stocks as described previously. In order to confirm the integrity of the SDM_*msDPCK* gene inserts, including the presence of the correct SDM mutated sites, sequencing PCR was performed on the respective isolated plasmids as described above.

4.3.4.2 Transformation of supplementary plasmids for co-expression

For the expression of the modified MBP-6xHis-tag fusion protein and the expression assisted by molecular chaperones, the transformation of an additional plasmid was required. This was achieved as follows:

Modified MBP-6xHis-tag (pRK586 supplementary plasmid) – Chemical competent cells were made (as described above) using the Invitrogen™ *E. coli* BL21 Star™ (DE3) cells that already contained the pMALcHT-*msDPCK* plasmid, followed by the transformation of the

supplementary plasmid (pRK586) into these chemical competent cells as previously described. However, these transformed cells were plated on LB agar plates supplemented with two antibiotics. The pRK586 plasmid (supplied by Sean T. Prigge as described in Kapust *et al.* [243]) is kanamycin resistant and expresses TEV protease for *in vivo* cleaving of the MBP-domain. Therefore, the LB agar plate was supplemented with ampicillin (100 mg/l) for the pMALcHT-*msDPCK* plasmid and kanamycin (30 mg/l) for the pRK586 plasmid.

Molecular chaperones (pG-KJE8 supplementary plasmid) – As described before, chemical competent cells were prepared using the Invitrogen™ *E. coli* BL21 Star™ (DE3) cells that already contain the SDM_pET28a(+)-*msDPCK*n plasmid. This was followed by the transformation of the relevant supplementary plasmid (pG-KJE8) into these chemical competent cells as described above. The pG-KJE8 plasmid (TaKaRa Bio Inc., Japan) is chloramphenicol (Sigma-Aldrich) resistant and expresses the necessary molecular chaperones to assist with protein folding. The LB agar plate was, therefore, supplemented with kanamycin (30 mg/l) for the SDM_pET28a(+)-*msDPCK*n plasmid and chloramphenicol (20 mg/l) for the pG-KJE8 plasmid.

Colony PCR was performed, following the transformation of the supplementary plasmids, in order to confirm the presence of the SDM_*msDPCK* insert as previously described. It was assumed that the supplementary plasmids were present as the single plasmid colonies failed to grow on double antibiotic supplemented LB agar plates. Subsequently, selected positively confirmed colonies were used to prepare O/N cultures, which were then used for expression trials, as well as making freezer stocks as described previously.

4.3.5 Expression and purification of Ms02 DPCK

4.3.5.1 Expression trials

First, the general procedure that was implemented for each expression trial is described, which is then followed by the description of the details of the expression conditions of each of the various soluble expression attempts.

The general procedure entailed the making of an O/N culture as previously described using the appropriate plasmid freezer stock (20 µl) or LB agar plate colony of the transformed Invitrogen™ *E. coli* BL21 Star™ (DE3) cells. This served as a starter culture for overexpression of the Ms02 DPCK protein. The expression trials were performed as follows: 5 ml growth media, supplemented by the relevant antibiotics, was added to a 15 ml Falcon® tube, followed by inoculation with 50 µl of starter culture under sterile conditions. The inoculated samples were incubated at 37˚C on an orbital shaker (150 rpm) until an average

optical density, measured at a wavelength of 600 nm (OD_{600}), of 0.6-0.7 was reached. Each sample was then induced with a predetermined concentration of isopropyl-β-D-1-thiogalactopyranoside (IPTG), followed by incubation at the appropriate temperature for a selected time.

For the expression trials where auto-induction medium (prepared according to the ZYM-5052 recipe as reported by Studier [253]) was used, the following procedure was followed: 15 ml of the prepared auto-induction medium, supplemented by the relevant antibiotics, was added to a 50 ml Erlenmeyer flask, followed by inoculation of the relevant starter culture and O/N incubation at 37˚C. Additionally, a 2 ml sample of the incubated culture was taken at 5 h.

The Novagen® BugBuster® Protein Extraction Reagent (Merck, RSA) was used, according to the manufacturer's instructions, to obtain the soluble fractions following the expression trials. These samples were analysed by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) as described in the following sections. In cases where the presence of a soluble protein fraction on the SDS-PAGE analysis was suspected, Western-blot or Dot-blot analyses were also performed, also described in the following sections. Various attempts at soluble expression of the target protein were made by testing a variety of expression conditions. These variable conditions that were tested in the expression trials for each plasmid set are described under the relevant subheadings below.

Expression via SDM pET28a(+)-msDPCK*n (N-terminal 6xHis-tag)* – Expression trials were performed using LB medium, Terrific Broth (TB) medium (12 g/l Bacto-Tryptone, 24 g/l yeast extract, 4 ml/l of 2 M glycerol) supplemented by 100 ml/l phosphate buffer (1 M KH_2PO_2, 1 M K_2HPO_4), and Auto-induction medium. The tested induction temperatures were 18˚C, 25˚C and 37˚C with varying IPTG concentrations of 100 µM, 200 µM and 400 µM. Furthermore, the incubation times after induction were 1 h and 2 h tested at 37˚C, as well as 4 h and O/N tested at all temperatures.

Expression via pET28a(+)-msDPCK*c (C-terminal 6xHis-tag)* – Expression trials were performed using LB medium, TB medium supplemented by phosphate buffer, and Auto-induction medium. The tested induction temperature was 37˚C with varying IPTG concentrations of 200 µM, 400 µM, 600 µM and 800 µM. The tested incubation intervals after induction were 1 h and 2 h, 4 h and O/N. Expression via pET28a(+)-*msDPCK*n was used as a positive control.

Expression via pMALcHT-msDPCK *(Modified MBP-6xHis-tag) and pRK586* – Since both plasmids are IPTG induced, the expression trials were performed using LB medium and

Auto-induction medium. Initially, O/N expression was tested with both plasmids at 20˚C and 37˚C with varying IPTG concentrations of 200 µM, 400 µM and 600 µM.

Expression trials via pMALcHT-*msDPCK* (without pRK586) were also performed under the same conditions. Additionally, following the suggested guidelines on improving solubility of MBP-tag fusion proteins by Waugh [241], alternative expression trials via pMALcHT-*msDPCK* (without pRK586) were performed. This entailed supplementing 10 ml LB medium with ampicillin (100 mg/l), which was then added to a 50 ml Greiner centrifuge tube, followed by inoculation with 100 µl of the appropriate starter culture. The samples were incubated as described previously and then induced with IPTG. The tested induction temperatures were 18˚C and 30˚C with varying IPTG concentrations of 10 µM, 50 µM and 100 µM, which were incubated for 18 h, 24 h and 48 h.

Expression via SDM_pET28a(+)-msDPCK*n and pG-KJE8 (molecular chaperones)* – Expression trials were performed using LB medium. However, these two plasmids are induced with different compounds. For chaperone expression by the pG-KJE8 plasmid, induction by arabinose and tetracycline is required, which is added to the LB medium before inoculation. These compounds add additional variables into the expression trials. Initially, 5 ml LB media was supplemented by the appropriate antibiotics, as well as arabinose (2 mg/ml) and tetracycline (10 ng/ml), followed by inoculation with 50 µl of the relevant starter culture. After the appropriate OD_{600} was reached (as described above), IPTG induction of 12 h was performed at 20˚C and 37˚C with varying concentrations of 200 µM, 400 µM and 600 µM.

The expression conditions, as described by Wipperman *et al.* [249], was also tested. Subsequently, IPTG induction concentrations of 100 µM, 400 µM and 600 µM were tested at temperatures of 18˚C and 25˚C for 20 h and 12 h, respectively. The arabinose (2 mg/ml) and tetracycline (10 ng/ml) concentrations were kept constant throughout these expression trials.

Additionally, five tubes containing 5 ml LB medium supplemented with the appropriate antibiotics, each with different combinations of plasmid (i.e. only SDM_pET28a(+)-*msDPCK*n or SDM_pET28a(+)-*msDPCK*n and pG-KJE8) and induction compounds (i.e. IPTG only, arabinose and tetracycline only or combination of both), were inoculated for expression trials as follows: (*i*) only SDM_pET28a(+)-*msDPCK*n with IPTG (400 µM) induction (positive control), (*ii*) only SDM_pET28a(+)-*msDPCK*n and no IPTG induction (negative-positive control), (*iii*) both plasmids with no arabinose, tetracycline with no IPTG induction (negative control), (*iv*) both plasmids with arabinose (0.5 mg/ml, as proposed by manufacturer's instructions) and tetracycline (5 ng/ml, as proposed by manufacturer's instructions) with no

IPTG induction (expression negative control), and (*v*) both plasmids with arabinose (0.5 mg/ml) and tetracycline (5 ng/ml) with IPTG (400 µM) induction. These trials were performed at IPTG induction intervals of 1 h, 2 h and O/N at 37˚C.

4.3.5.2 SDS-PAGE analysis

The SDS-PAGE analyses were performed with a 4.5% acrylamide stacking gel and a 12% acrylamide resolving gel. The resolving gel was cast by adding 4.8 ml acrylamide stock (30% T, 2.7% C) to 7.2 ml resolving buffer (375 mM Tris-base, 0.1% (m/v) SDS, pH 8), followed by the addition of 30 µl of N,N,N',N'-tetramethylethylenediamine (TEMED, Sigma-Aldrich) and 36 µl of ammonium persulphate (200 µg/µl). The stacking gel was casted by adding 0.75 ml acrylamide stock (30% T, 2.7% C) to 4.25 ml stacking buffer (125 mM Tris-base, 0.1% (m/v) SDS, pH 6.8), followed by the addition of 15 µl of TEMED and 30 µl of ammonium persulphate (200 µg/µl).

Before electrophoresis, the samples were prepared for loading as follows: 30 µl of soluble fraction was added to a 1.5 ml microcentrifuge tube, followed by addition of 30 µl of reducing treatment buffer (125 mM Tris-base, 4% (m/v) SDS, 20% (v/v) glycerol, 10% (v/v) β-mercaptoethanol, pH 6.8) and 12 µl of bromophenol blue solution (0.1% (m/v) bromophenol blue in 100 mM NaOH). Similar mixtures were prepared for the insoluble fractions, but the insoluble fractions were resuspended in 200 µl of Milli-Q® water prior to preparation. These mixtures were then heat treated by boiling for 5 min, followed by placing the mixtures on ice until it was loaded (20 µl) for electrophoresis. Additionally, 6 µl of Broad Range Color Prestained Protein Standard (New England Biolabs) was loaded as a molecular weight marker.

Electrophoresis was performed in an electrode buffer (25 mM Tris-base, 192 mM glycine, 0.1% SDS. pH 8.3) at a constant current of 25 mA for 1.5-2 h. Subsequently, the resolving gel was stained by Coomassie staining solution (0.125% (m/v) Coomassie Brilliant Blue R250, 50% (v/v) methanol, 10% (v/v) glacial acetic acid) for 1 h at 37˚C with agitation in a hybridisation oven/shaker. After decanting the staining solution, the gel was placed in destain solution 1 (50% (v/v) methanol, 10% (v/v) glacial acetic acid) for 1 h 30 min under agitation at 37˚C, during which the destain solution was replaced by fresh solution every 30 min. The destaining of the gel was completed by incubating it O/N in destain solution 2 (5% (v/v) methanol, 7% glacial acetic acid) at room temperature under agitation.

4.3.5.3 Western-blot and Dot-blot analyses

The SDS-PAGE analysis was performed as described above, but the Spectra™ Multicolor Broad Range Protein Ladder (Thermo Fisher Scientific) was used as a molecular weight

marker and the resolving gel was not stained. Instead, the resolving gel was used to transfer the proteins to a 0.45 µm nitrocellulose membrane (Schleicher & Schuell, Sigma-Aldrich) by electrophoresis (up to 1 A, 25 V) for 30 min via the Trans® Turbo™ Blotting System (Bio-Rad, RSA). The nitrocellulose membrane was, subsequently, placed in blocking buffer (25 mM Tris-HCl, 150 mM NaCl, 2% (m/v) bovine serum albumin (BSA), pH 7.6) for 1 h under agitation at room temperature to prevent non-specific binding. This was followed by washing the membrane twice for 15 min with 1x Tris-buffered saline solution containing Tween (TBST, 25 mM Tris-HCl, 150 mM NaCl, 0.05% Tween®-20, pH 7.6). The membrane was then treated with the HisProbe-HRP (Thermo Fisher Scientific) working solution (diluted 1:5 000 with blocking buffer) for 1 h under agitation at room temperature. Next, the membrane was washed four times with TBST for 10 min before visualisation. The membrane was developed for up to 30 min in the 1x Phosphate-buffered saline solution containing 4-chloro-1-naphthol, methanol and H_2O_2 (10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 137 mM NaCl, 2.7 mM KCl, 0.05% (m/v) 4-chloro-1-naphthol, 16% (v/v) methanol, 0.025% (v/v) H_2O_2).

The Dot-blot analysis is similar to the Western-blot analysis; however, no SDS-PAGE analysis is required. Instead, 1-2 µl of the protein sample is added directly onto the nitrocellulose membrane and is left to dry (completely). Thereafter, the same procedure as described for the Western-blot analysis is followed, continuing on from the blocking step.

4.3.5.4 Sarkosyl detergent treatment

Before treatment with the sarkosyl detergent, expression with the SDM_pET28a(+)-*msDPCK*n plasmid containing *E. coli* cells was performed at large-scale. This was done as follows: 500 ml LB medium, supplemented by kanamycin (30 mg/l), was added to a 2 l Erlenmeyer flask, followed by inoculation with 5 ml of the appropriate starter culture. The inoculated sample was incubated at 37˚C on an orbital shaker (150 rpm) until an average OD_{600} of 0.6-0.7 was reached. The cells were induced with 400 µM IPTG and incubated O/N at 18˚C on an orbital shaker at 150 rpm. Subsequently, the cells were harvested and lysed according to the protocol provided by Massiah *et al.* [252], with one deviation, i.e. the cells were gently mixed at a low speed on an orbital shaker (25 rpm) at 4˚C after the addition of sarkosyl (Sigma-Aldrich). This was to prevent the possible denaturing of the enzyme. The cell lysate (and pellet) was analysed by SDS-PAGE and Western-blot as described above.

4.3.5.5 Purification of large-scale sarkosyl treated expressions

Purification was performed using the ÄKTA*prime*-system. The cell lysate was first filtered using Acrodisc® 25 mm syringe filters with a pore size of 0.45 µm (Thermo Fisher Scientific) before proceeding to injection into the ÄKTA*prime*-system. The 6xHis-tagged target protein, along with binding buffer (20 mM Tris-HCl, 5 mM Imidazole, 500 mM NaCl, 0.05% (m/v)

Sodium aside, 1% sarkosyl, pH 7.9), was loaded onto a 1 ml HiTrap® Chelating High Performance column (Amersham Biosciences, USA) that was preloaded Ni^{2+}. An optimised wash step entailing of 20 mM imidazole was implemented to remove any non-specific binding. Optimisation was performed by varying the Imidazole concentration in this step. This was followed by the elution of the 6xHis-tagged target protein by increasing the imidazole concentration with the elution buffer (20 mM Tris-HCl, 500 mM Imidazole, 500 mM NaCl, 0.05% (m/v) Sodium aside, 1% sarkosyl, pH 7.9). The protein elution was monitored by ultraviolet at a wavelength of 280 nm. All purified samples were desalted using a 5 ml HiTrap® Desalting column (Amersham Biosciences, USA) and the buffer exchange buffer (25 mM Tris-HCl, 75 mM NaCl, pH 8) via the ÄKTA*prime*-system. Subsequently, glycerol was added to the samples (5% (v/v) final concentration), and stored at -20˚C.

The purified samples were analysed by SDS-PAGE and Western-blot as described above. Furthermore, a non-reducing SDS-PAGE analysis was performed. This was performed by using a non-reducing treatment buffer (125 mM Tris-base, 4% (m/v) SDS, 20% (v/v) glycerol, 20 mM Iodoacetamide, pH 6.8), instead of the described reducing treatment buffer.

4.3.5.6 Bradford protein concentration assay

The protein concentrations of the purified samples were determined using the Quick Start™ Bradford Protein Assay kit (Bio-Rad), which consisted of a set of BSA protein standards and Bradford reagent. A standard curve with a concentration range of 0.0-1.0 mg/ml was set up with BSA standards (diluted in Milli-Q® water) with concentrations of 0.0 (blank), 0.125, 0.25, 0.5, 0.75 and 1.0 mg/ml. This was performed in a flat-bottom 96-well polystyrene microtitre plate (Greiner Bio-one, Germany) by adding 5 µl of each BSA concentration standard per well, followed by the addition of 250 µl Bradford reagent (each standard concentration was done in triplicate). The plate was incubated at room temperature for 10-15 min in a dark environment. The absorbance at 595 nm was measured using a Varioskan microplate reader (Thermo Fisher Scientific). The standard curve was then used as a reference to determine the concentration of the purified samples by performing the same procedure as with the standards (also done in triplicate). If the concentration of the samples were above the range of the standard curve, the sample was diluted and the assay repeated accordingly.

4.3.6 Preliminary test for activity of Ms02 DPCK

The enzymatic activity of the purified Ms02 DPCK enzyme was determined by performing *in vitro* enzyme reactions, followed by analysis via HPLC. The 50 µl reaction mixtures were incubated for 1 h at 37˚C and consisted of the following: 50 mM Tris-HCl buffer (pH 7.6), 20 mM KCl, 10 mM $MgCl_2$, 1 mM Tris(2-carboxyethyl)phosphine, 1.8 mM phosphoenol-

pyruvate, 0.04 U/µl pyruvate kinase, 1.5 mM ATP, 0.25 mM DePCoA and 0.1 mg/ml DPCK enzyme. Subsequently, the derivatisation and HPLC analysis were performed as described by Goosen *et al*. [231]. The purified DPCK enzyme of *E. coli* was used as positive control.

4.4 Results

4.4.1 Amplification, cloning and site-directed mutagenesis of *msDPCK*

4.4.1.1 Amplification and isolation of msDPCK *using Ms02 genomic DNA*

The *msDPCK* gene (579 bp) could be successfully amplified at three of the five tested annealing temperatures (50˚C, 55˚C, 60˚C) by using Ms02 genomic DNA as template, which can be observed on the agarose gel analysis of the PCR products (Figure 4.4 A). Additionally, the agarose gel analysis showed that each of the annealing temperatures produced a single PCR product with no non-specific binding.

Consequently, the T_a of 55˚C was selected for large-scale amplification and subsequent purification as the PCR product at this T_a produced a thick and bright band at the expected size (~600 bp), which indicated a high concentration of product. A T_a of 50˚C also produced a thick and bright band; however, the selected T_a of 55˚C was closer to the estimated T_m of ±58˚C (according to the IDT specification sheet) and was, therefore, the preferred choice. The subsequent pooled PCR products were successfully purified producing a single band at the expected size (~600 bp, Figure 4.4 B). The estimated concentration of the purified *msDPCK* PCR product was determined to be 273.7 ng/µl.

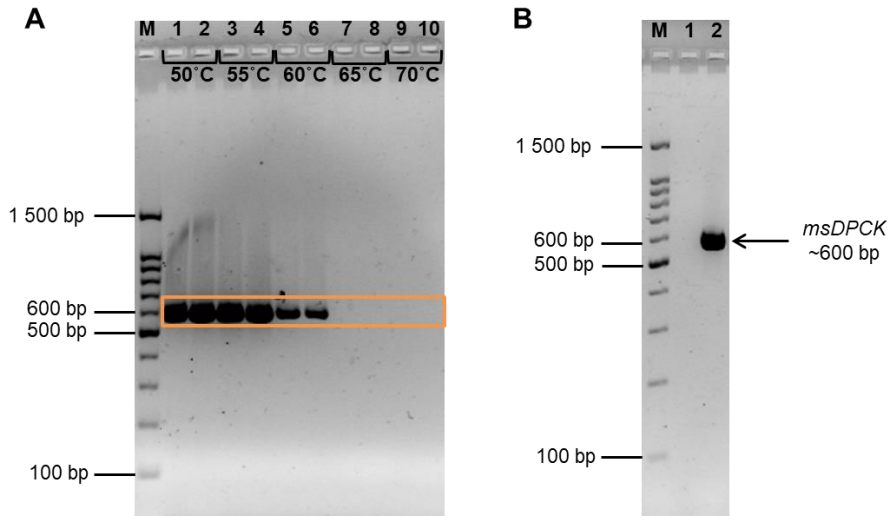


Figure 4.4 2% Agarose gels showing the PCR product of the amplification of the *msDPCK* gene using the Ms02 genomic DNA as template. (A) Optimisation of annealing temperatures (T_a). The orange box indicates the expected position of the amplified product. M: 100 bp DNA ladder marker, Lane 1-2: T_a = 50˚C, Lane 3-4: T_a = 55˚C, Lane 5-6: T_a = 60˚C, Lane 7-8: T_a = 65˚C, Lane 9-10: T_a = 70˚C. (B) Pooled and purified PCR product. M: 100 bp DNA ladder marker, Lane 1: Open, Lane 2: Isolated *msDPCK* gene with a size of about 600 bp.

4.4.1.2 Cloning of msDPCK *PCR product into pET28a(+) with an N-terminal 6xHis-tag for subsequent SDM*

Colony PCR was performed (using Primer set 1 in Table 4.1) in order to analyse the *E. coli* JM109 colonies that formed after transformation with the ligation reaction of the RE digested products (digested with the NdeI and XhoI REs). The agarose gel analysis of the colony PCR products (Figure 4.5) confirmed the presence of the *msDPCK* gene insert within 14 of the 17 tested colonies. The positively confirmed colonies can be identified by the presence of a single bright band at the expected size (~600 bp). Conversely, the colonies that contained no insert (Lane 1, 2 and 5 in Figure 4.5) could be identified by the lack of a PCR product band.

The presence of the *msDPCK* gene insert within the positively confirmed colonies, thus, suggests that the RE digestion, ligation and subsequent transformation of the pET28a(+)-*msDPCK*n plasmid were successful. Furthermore, the colony PCR negative control (Lane 18 in Figure 4.5) did not result in any PCR product band, which further strengthens the reliability of the colony PCR analysis.

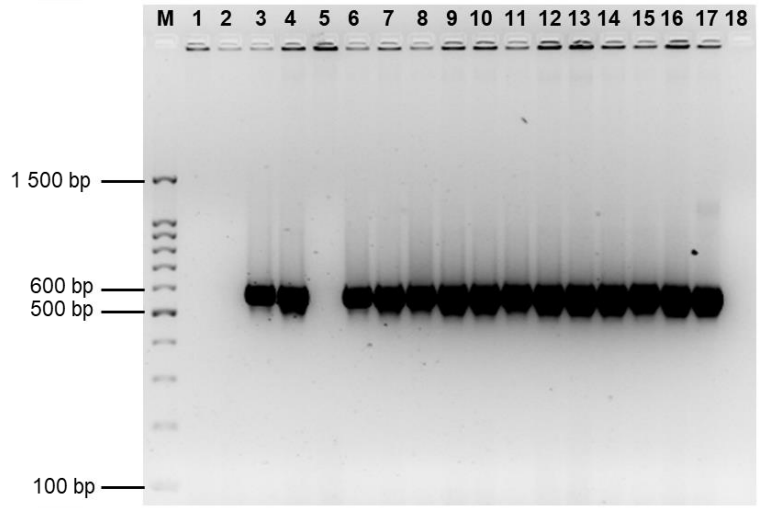


Figure 4.5 2% Agarose gel of the colony PCR products (using Primer set 1) after RE digestion, ligation and transformation of the pET28a(+)-*msDPCK*n plasmid into *E. coli* JM109 cells. M: 100 bp DNA ladder marker, Lane 1-17: PCR products of the colonies tested whereof the colonies of Lane 1, 2 and 5 tested negative for the *msDPCK* gene insert, Lane 18: Negative control.

4.4.1.3 Site-directed mutagenesis of msDPCK

After digestion with DpnI, the mutated plasmid product of the SDM PCR was transformed into *E. coli* JM109 cells, and analysed by colony PCR (results not shown). Six positively tested colonies were selected for O/N cultures and the plasmid isolations of these cultures were, subsequently, analysed by sequencing PCR. Of the six analysed plasmid inserts, two

inserts were identified to have the desired mutated sites while retaining an otherwise correct and complete *msDPCK* gene sequence.

The full sequence alignment of the *msDPCK* gene and the sequenced inserts of the two plasmid isolations (Supplementary Figure 2.1) is shown in Appendix 2, whereas a condensed version showing segments of this sequence alignment is depicted in Figure 4.6. The desired mutated sites can be observed at position 561 and 567 in the sequence alignment, which indicates that the TGA codon (red box in Figure 4.6) was mutated to produce a TGG codon. Thus, the SDM PCR was successful in these two particular plasmid isolations.

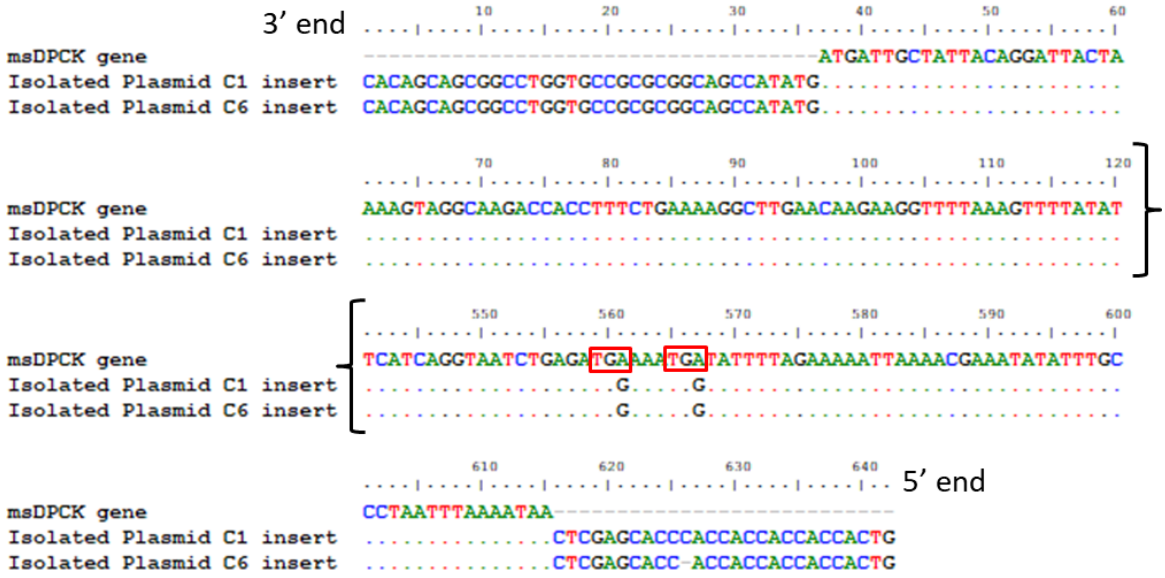


Figure 4.6 A condensed sequence alignment showing segments of the sequencing results subsequent to the SDM of the *msDPCK* gene insert. The *msDPCK* gene, along with the sequenced inserts of the isolated plasmids from colony 1 and 6 are shown, respectively. The red boxes indicate the TGA codons in the *msDPCK* gene. The black brackets indicate the start and end position of the section of the sequence that is not shown.

Subsequently, both of these isolated plasmids were subjected to RE digestion with NdeI and XhoI, followed by the isolation of the SDM_*msDPCK* gene inserts. The respective isolated inserts were then ligated into new pET28a(+) vectors and the resulting plasmids transformed into *E. coli* JM109 cells, which was analysed by colony PCR using primers of the T7 primer set (Figure 4.7).

Accordingly, only half of the 12 tested colonies were confirmed positive for the correct insert, which can be seen by the single colony PCR product band (~750 bp) in Lane 3, 4, 5, 6, 9 and 12 of the agarose gel analysis (Figure 4.7). The amplified product is slightly larger in size since the primers of the T7 primer set (Table 4.3) was used; thus, amplifying part of the

pET28a(+) vector, along with the SDM_*msDPCK* gene insert. This suggests that the RE digestion, ligation and succeeding transformation of the SDM_pET28a(+)-*msDPCK*n plasmid were successful within these colonies. Moreover, the reliability of the colony PCR is strengthened by the lack of any colony PCR product band in the negative control lane of the agarose gel analysis (Lane 13).

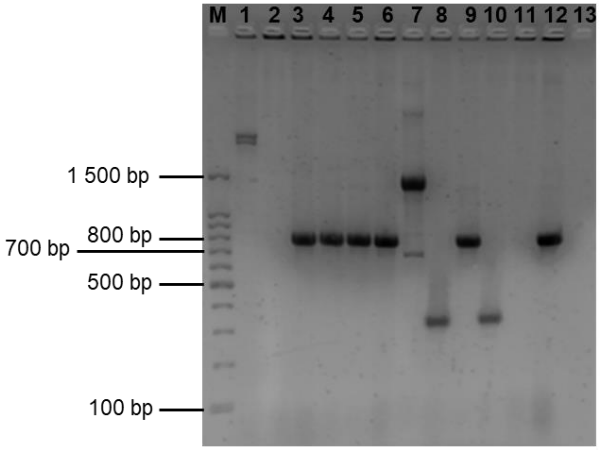


Figure 4.7 2% Agarose gel of the colony PCR products (using the T7 primer set) after RE digestion, ligation and transformation of the SDM_pET28a(+)-*msDPCK*n plasmid into *E. coli* JM109 cells. M: 100 bp DNA ladder marker, Lane 1-12: PCR products of the colonies tested whereof the colonies of Lane 3, 4, 5, 6, 9 and 12 tested positive for the correct *msDPCK* gene insert, Lane 13: Negative control.

Two of the six colonies that were confirmed to be positive for the correct SDM_*msDPCK* gene insert were used to make O/N cultures for the purpose of plasmid isolation. These respective plasmid isolations were then transformed into Invitrogen™ *E. coli* BL21 Star™ (DE3) cells, which was analysed by colony PCR. Overnight cultures of three positively confirmed colonies were made, followed by plasmid isolations, which were used for sequencing PCR. The sequence alignment of the sequencing PCR results (Supplementary Figure 2.2) provided the final confirmation of a successful SDM. The SDM_pET28a(+)-*msDPCK*n plasmid isolated from the third colony was, subsequently, used in following experiments since it displayed no sequence variation in its pET28a(+) vector component or its *msDPCK* gene insert component (apart from the desired SDM mutations).

4.4.2 Amplification and sub-cloning of SDM_*msDPCK*

*4.4.2.1 C-terminal 6xHis-tag plasmid (pET28a(+)-*msDPCK*c)*

The T_a optimisation PCR revealed an optimal T_a of 50˚C for the amplification of the SDM_*msDPCK* gene insert using the SDM_pET28a(+)-*msDPCK*n plasmid as template. The SDM_*msDPCK* gene insert was successfully amplified and purified (estimated concentration of 357.3 ng/µl), which was succeeded by RE digestion (using XhoI and NcoI), ligation and

transformation (into *E. coli* JM109 cells). The subsequent agarose gel analysis of the colony PCR products (Figure 4.8) confirmed the presence of the SDM_*msDPCK* gene insert in all of the tested colonies, which can be seen by the single amplified PCR product band (~600 bp). The negative control of the PCR also strengthened the reliability of the colony PCR since no product band can be observed in Lane 11.

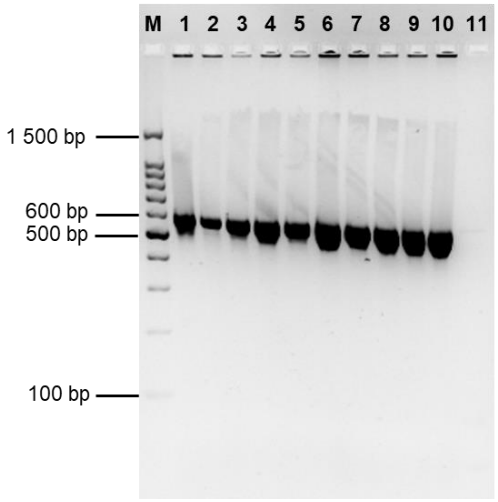


Figure 4.8 2% Agarose gel of the colony PCR products (using Primer set 2) after RE digestion, ligation and transformation of the pET28a(+)-*msDPCK*c plasmid into *E. coli* JM109 cells. M: 100 bp DNA ladder marker, Lane 1-10: PCR products of the tested colonies where all colonies tested positive for the SDM_*msDPCK* gene insert, Lane 11: Negative control.

*4.4.2.2 Modified MBP-6xHis-tag plasmid (pMALcHT-*msDPCK*)*

The PCR amplification of the SDM_*msDPCK* gene insert using the SDM_pET28a(+)-*msDPCK*n plasmid as template was successfully performed at a T_a of 50˚C, which was determined to be the optimal T_a by PCR optimisation tests. The purity of pooled and purified PCR amplification products was confirmed by agarose gel analysis (Lane 3 in Figure 4.9 A) and had an estimated concentration of 301.2 ng/µl. The products of the RE digestion (using EcoRI and SalI) was analysed before ligation (Figure 4.9 A). The isolated pSPr022 plasmid (Lane 1) was digested, the *P. falciparum* ACP gene insert removed and the resulting plasmid (pMALcHT) was loaded in Lane 2. A single product band of the pMALcHT plasmid (orange box in Lane 2) can be seen at a smaller size, compared to that of the isolated pSPr022 plasmid and, therefore, suggests the removal the insert. Moreover, the single product band indicates that the pMALcHT plasmid is pure. Similarly, this can be observed for the RE digestion product of the purified PCR product (red box in Lane 4). Thus, the RE digestion was successful for both the amplified SDM_*msDPCK* gene insert and the pSPr022 plasmid.

Subsequently, the successive ligation and transformation was performed and analysed by colony PCR (using the primers of Primer set 3). All of the tested colonies were confirmed to

be positive for the SDM_*msDPCK* gene insert, which can be seen by the single product band (Lane 1-6) between 500 bp and 600 bp according to the marker (Figure 4.9 B). The lack of a product band in the negative control (Lane 7) also gave added support for the reliability of the colony PCR. Additionally, the inserts of three positively confirmed colony plasmid isolations were sequenced. This provided additional confirmation of the correct gene insert, including the presence of the desired TGG mutations (results not shown).

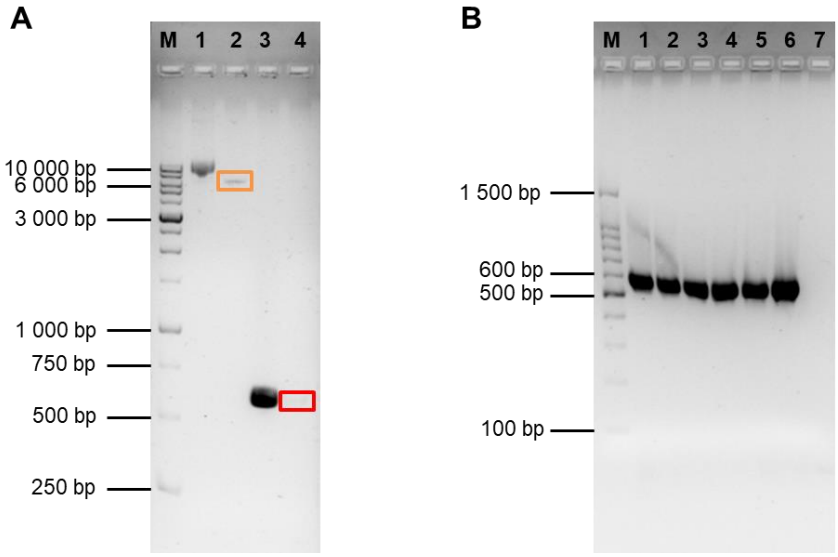


Figure 4.9 Agarose gels of the analyses subsequent to the amplification of the SDM_*msDPCK* gene insert (using Primer set 3). (A) 1% Agarose analysis of the RE digestion products before ligation. The orange box indicates the pMALcHT plasmid (pSPr022 after digestion and removal of *P. falciparum* ACP gene insert). The red box indicates the SDM_*msDPCK* gene insert. M: 1 kb DNA ladder marker, Lane 1: Isolated pSPr022 plasmid, Lane 2: pMALcHT plasmid after dephosphorylation, Lane 3: Pooled and purified PCR product. Lane 4: Purified PCR product after RE digestion. (B) 2% Agarose analysis of the colony PCR products (using Primer set 3) after ligation and transformation of the pMALcHT-*msDPCK* plasmid into *E. coli* JM109 cells. M: 100 bp DNA ladder marker, Lane 1-6: PCR products of the tested colonies where all colonies tested positive for the SDM_*msDPCK* gene insert, Lane 7: Negative control.

4.4.3 Transformation of plasmids prior to expression

*4.4.3.1 N-terminal 6xHis-tag plasmid (SDM_pET28a(+)-*msDPCK*n)*

The results of the transformation of the SDM_pET28a(+)-*msDPCK*n plasmid into Invitrogen™ *E. coli* BL21 Star™ (DE3) cells is reported in Section 4.4.1.3. This transformation was successful, which was confirmed by the colony PCR, as well as sequencing PCR (Supplementary Figure 2.2).

*4.4.3.2 C-terminal 6xHis-tag plasmid (pET28a(+)-*msDPCK*c)*

The isolated pET28a(+)-*msDPCK*c plasmid was successfully transformed into Invitrogen™ *E. coli* BL21 Star™ (DE3) cells. This was confirmed by colony PCR, as well as PCR

sequencing of the isolated plasmid inserts of two positively tested colonies using the forward and reverse primer of the T7 primers for the respective colonies. The sequence alignment of the sequencing results (Figure 4.10) indicates that the inserts of both selected colonies are correct and also properly ligated into the pET28a(+) vector with no stop codon (red box), which is as expected since the 6xHis-tag is at the C-terminal end.

*4.4.3.3 Modified MBP-6xHis-tag plasmids (pMALcHT-*msDPCK *and pRK586)*

The successful transformation of the pMALcHT-*msDPCK* plasmid into Invitrogen™ *E. coli* BL21 Star™ (DE3) cells was confirmed by colony PCR using Primer set 3 (Table 4.1). All of the colonies that were tested exhibited a single PCR product at the expected size (Figure 4.11 A). Subsequently, these *E. coli* cells were transformed with the pRK586 plasmid, which was analysed by colony PCR using the same primers as before. Even though these primers only confirmed the presence of the SDM_*msDPCK* gene insert (Figure 4.11 B), it was assumed that the pRK586 plasmid is also present in all of the positively tested colonies since the transformed cells were plated on a LB agar plate supplemented by two different antibiotics. Therefore, in order for the colonies to grow, both plasmids had to be successfully transformed to provide the necessary antibiotic resistance.

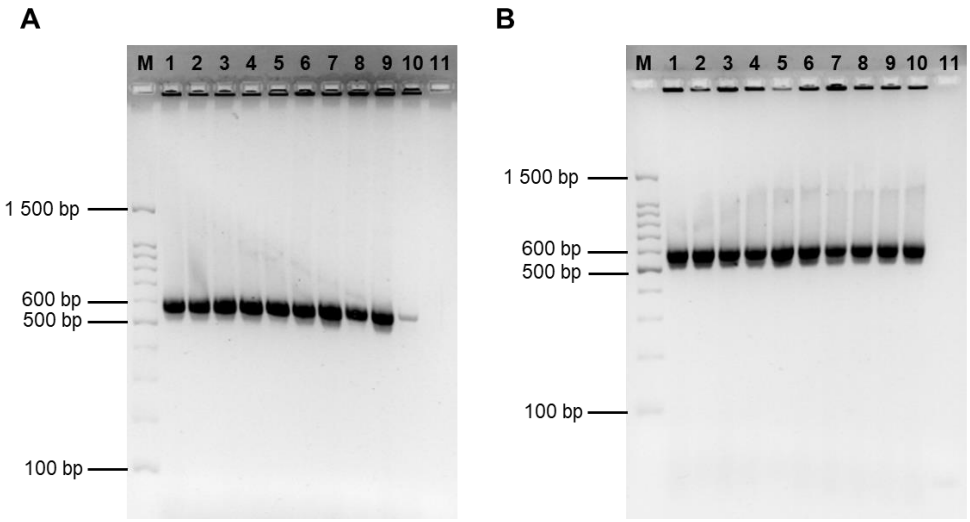


Figure 4.11 2% Agarose gels of the colony PCR products (using Primer set 3) after transformation into Invitrogen™ *E. coli* BL21 Star™ (DE3) cells. M: 100 bp DNA ladder marker, Lane 1-10: PCR products of the tested colonies where all colonies tested positive for the SDM_*msDPCK* gene insert, Lane 11: Negative control. (A) Colony PCR following transformation of pMALcHT-*msDPCK* (single plasmid). (B) Colony PCR following transformation of pRK586 (double plasmid).

Figure 4.10 Sequence alignment of the PCR sequencing results following the transformation of the pET28a(+)-*msDPCK*c plasmid into Invitrogen™ *E. coli* BL21 Star™ (DE3) cells. The red box indicates the expected lack of a stop codon in the sequenced plasmid inserts. The black box indicates the 6xHis-tag.

4.4.3.4 Molecular chaperone expressing plasmid (pG-JKE8)

The pG-JKE8 plasmid was transformed into Invitrogen™ *E. coli* BL21 Star™ (DE3) cells that consisted of the previously transformed SDM_pET28a(+)-*msDPCK*n plasmid. The colony PCR analysis (Figure 4.12) only confirmed the presence of the SDM_*msDPCK* gene insert in the tested colonies. The transformation of the pG-JKE8 plasmid was, however, presumed to be successful due to the growth of positively tested colonies on an LB agar plate supplemented with two different antibiotics.

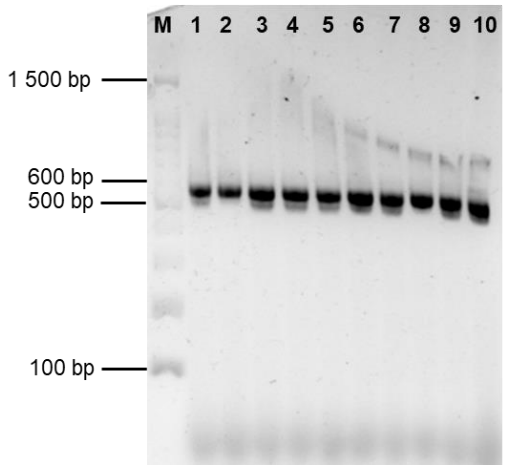


Figure 4.12 2% Agarose gels of the colony PCR products (using Primer set 1) after transformation of the pG-JKE8 plasmid into Invitrogen™ *E. coli* BL21 Star™ (DE3) cells containing the SDM_pET28a(+)-*msDPCK*n plasmid. M: 100 bp DNA ladder marker, Lane 1-10: PCR products of the tested colonies where all colonies tested positive for the SDM_*msDPCK* gene insert.

4.4.4 Attempts at soluble expression of Ms02 DPCK

*4.4.4.1 N-terminal 6xHis-tag plasmid (SDM_pET28a(+)-*msDPCK*n)*

Prior to the expression trials, the Ms02 DPCK recombinant protein (MsDPCK, 22.7 kDa) with an N-terminal 6xHis-tag was initially expressed on a large scale at 37˚C for 2 h and 4 h with 400 µM IPTG followed by the purification of the soluble protein fractions. The SDS-PAGE analysis of these large-scale expressions and subsequent purifications (Figure 4.13 A) revealed that the expressed product is entirely in the insoluble protein fraction since there was no visible protein band at the expected size in the purified samples (Lane 3 and Lane 4, Figure 4.13 A). Additional confirmation of this was provided by the Western-blot analysis (Figure 4.13 B), which indicated a positive signal for the 6xHis-tag protein at the expected size (±24 kDa) in the insoluble protein fractions (orange box) but not in the purified samples. Consequently, various expression trials were initiated to attempt soluble expression of the MsDPCK protein.

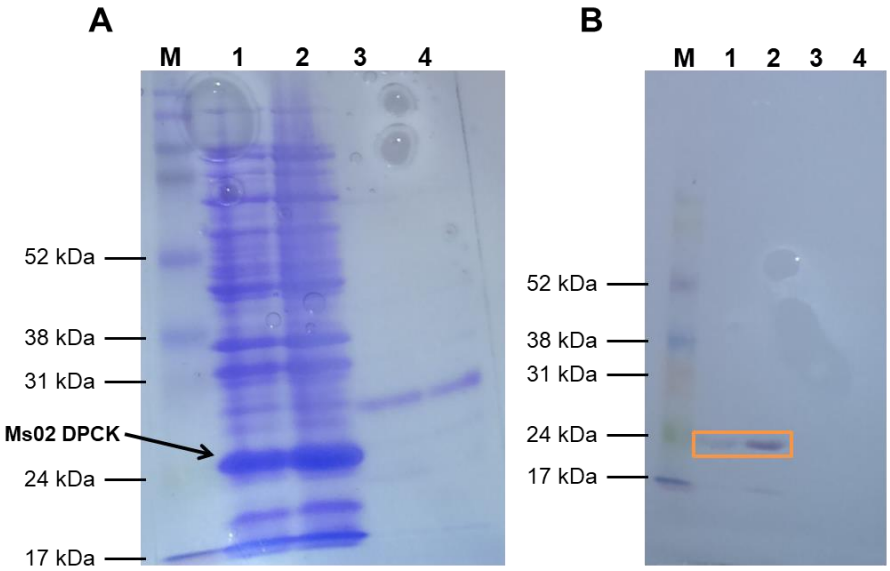


Figure 4.13 Analysis of initial large-scale expression of MsDPCK with an N-terminal 6xHis-tag at 37˚C for 2 h and 4 h following purification. M: Amersham™ Full-range Rainbow™ marker (Sigma-Aldrich), Lane 1: Insoluble protein fraction of 2 h expression, Lane 2: Insoluble protein fraction of 4 h expression, Lane 3: Purified sample of the soluble protein fraction of 2 h expression, Lane 4: Purified sample of the soluble protein fraction of 4 h expression. (A) 12% SDS-PAGE analysis of insoluble protein fractions and purified factions. The MsDPCK protein product band can be seen at ±24 kDa. (B) Corresponding Western-blot analysis of insoluble protein fractions and purified factions. The orange box indicates the positive signal for the 6xHis-tag in the insoluble protein fractions.

The O/N expression trials of the N-terminal 6xHis-tagged MsDPCK protein (MsDPCK_N-His) at 18˚C, 25˚C and 37˚C with varying IPTG induction concentrations all failed to produce a soluble target protein. This was confirmed by the SDS-PAGE analysis of the soluble and insoluble protein fractions of these expression trials (Figure 4.14).

In the soluble protein fractions (Figure 4.14 A), no clear band difference between the uninduced samples (Lane 1 and Lane 5) and the relevant induced samples could be observed. On the other hand, the SDS-PAGE analysis of the insoluble fractions (Figure 4.14 B) showed that overexpression of the target protein did indeed occur at all three temperatures (black box). This, thus, suggested that no soluble expression of the target protein occurred and was further confirmed by Western-blot (results not shown). Similar results were obtained under the various conditions stated in Section 4.3.5.1 for the expression trials using the SDM_pET28a(+)-*msDPCK*n plasmid.

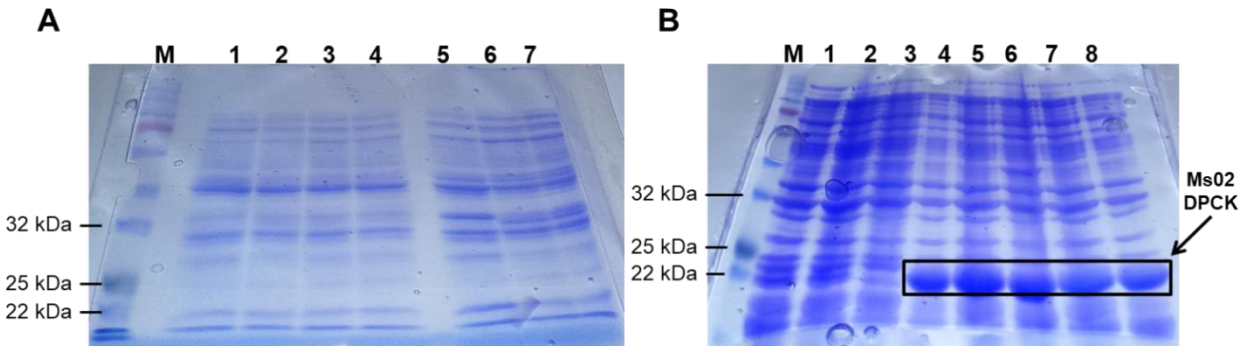


Figure 4.14 12% SDS-PAGE analysis of O/N expression trials of MsDPCK_N-His at 18˚C, 25˚C and 37˚C. M: Broad Range Protein Standard marker. (A) Soluble protein fractions at 18˚C and 37˚C. Lane 1: 0 µM IPTG at 18˚C (uninduced sample), Lane 2: 100 µM IPTG at 18˚C, Lane 3: 200 µM IPTG at 18˚C, Lane 4: 400 µM IPTG at 18˚C, Lane 5: 0 µM IPTG at 37˚C (uninduced sample), Lane 6: 100 µM IPTG at 37˚C, Lane 7: 200 µM IPTG at 37˚C. (B) Insoluble protein fractions at 18˚C, 25˚C and 37˚C. Lane 1: 0 µM IPTG at 18˚C (uninduced sample), Lane 2: 0 µM IPTG at 25˚C (uninduced sample), Lane 3: 0 µM IPTG at 37˚C (uninduced sample), Lane 4: 200 µM IPTG at 18˚C, Lane 5: 200 µM IPTG at 25˚C, Lane 6: 200 µM IPTG at 37˚C, Lane 7: 400 µM IPTG at 18˚C, Lane 8: 400 µM IPTG at 25˚C. Black box indicates overexpression of MsDPCK_N-His.

*4.4.4.2 C-terminal 6xHis-tag plasmid (pET28a(+)-*msDPCK*c)*

The expression of the C-terminal 6xHis-tagged MsDPCK protein (MsDPCK_C-His) proved to be very problematic as no overexpression of the target protein (soluble or insoluble) could be achieved. This recurring result was obtained after all of the expression trials under the conditions described in Section 4.3.5.1 using the pET28a(+)-*msDPCK*c plasmid.

As an example, the SDS-PAGE analysis of O/N expression trials at 37˚C with varying IPTG concentrations is shown (Figure 4.15). The expression of the MsDPCK_N-His protein was included in the trials as a positive expression control. The results of the SDS-PAGE analysis of the soluble protein fractions (Figure 4.15 A) revealed the lack of soluble overexpression of the target protein for both MsDPCK_N-His and MsDPCK_C-His since there are no observable differences between the uninduced (Lane 1 and Lane 3, respectively) and induced samples. Similar results were observed for the corresponding insoluble protein fractions of the MsDPCK_C-His expression trials (Lane 3-6, Figure 4.15 B). However, the induced positive control (Lane 2, Figure 4.15 B) displayed overexpression of the MsDPCK_N-His protein (indicated by the arrow), which suggests that the expression conditions were correctly executed. The results of the SDS-PAGE analysis were also confirmed by a subsequent Western-blot analysis (not shown). Therefore, under the conditions that were tested in this study, the MsDPCK_C-His protein could not be overexpressed.

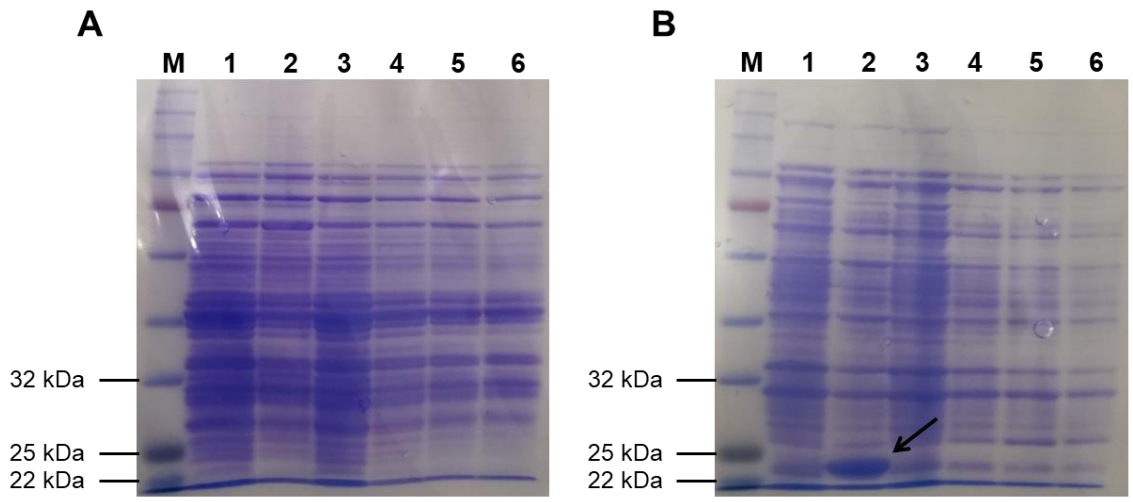


Figure 4.15 12% SDS-PAGE analysis of O/N expression trials of MsDPCK_C-His at 37˚C with the expression of MsDPCK_N-His as a positive expression control. M: Broad Range Protein Standard marker, Lane 1: 0 µM IPTG for MsDPCK_N-His (uninduced positive control), Lane 2: 400 µM IPTG for MsDPCK_N-His (induced positive control), Lane 3: 0 µM IPTG for MsDPCK_C-His (uninduced sample), Lane 4: 200 µM IPTG for MsDPCK_C-His, Lane 5: 400 µM IPTG for MsDPCK_C-His, Lane 6: 800 µM IPTG for MsDPCK_C-His. (A) Soluble protein fractions. (B) Insoluble protein fractions. The arrow indicates to the insoluble overexpressed MsDPCK_N-His protein band.

*4.4.4.3 Modified MBP-6xHis-tag plasmids (pMALcHT-*msDPCK *and pRK586)*

The co-expression of the pMALcHT-*msDPCK* and pRK586 plasmids, which produces the MBP-6xHis-tagged MsDPCK protein (MsDPCK_MBP-His) and TEV protease respectively, also failed to result in the overexpression of a soluble target protein. This can be seen in the SDS-PAGE analysis of the O/N expression trials using these plasmids with differing IPTG concentrations at 20˚C and 37˚C (Figure 4.16). In the soluble protein fractions of the expression trials (Figure 4.16 A), there seemed to be no clear overexpression of the target protein at both temperatures or with any of the selected IPTG concentrations. However, an overexpressed protein band (indicated by the arrow) could be observed in the three induced samples at 37˚C (Lane 6-7, Figure 4.16 A). This band is slightly below the 46 kDa marker band, which suggests that this protein band might be MBP (42 kDa). If this was the case, it would imply that the MsDPCK_MBP-His protein was expressed and cleaved successfully by the TEV protease. Nevertheless, the cleaved target protein was still not observed in the soluble protein fraction.

The SDS-PAGE analysis of the insoluble protein fractions (Figure 4.16 B) revealed bands that suggested that overexpression of the target protein did possibly occur. The same result was observed after expression trials with Auto-induction medium. However, this could not be confirmed using Western-blot analysis, as no signal for the His-tag could be detected.

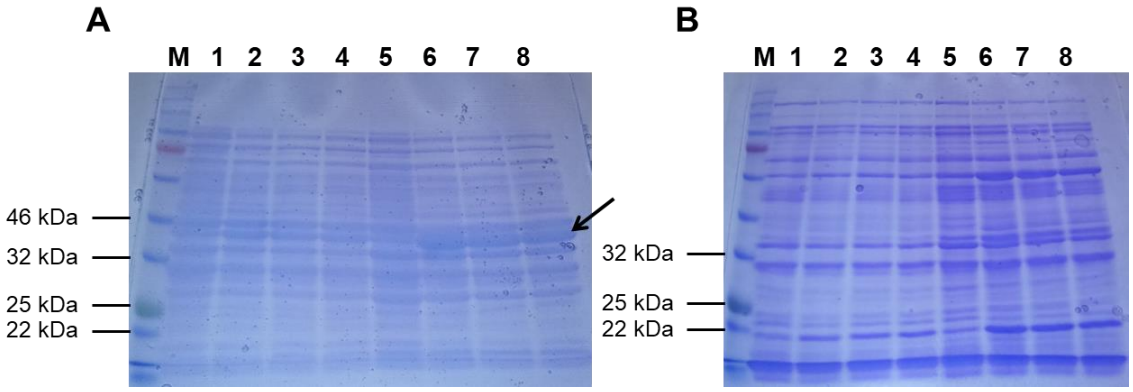


Figure 4.16 12% SDS-PAGE analysis of O/N expression trials at 20˚C and 37˚C using the double plasmid containing *E. coli* cells for expression of MsDPCK_MBP-His and TEV protease. M: Broad Range Protein Standard marker, Lane 1: 0 µM IPTG at 20˚C (uninduced sample), Lane 2: 200 µM IPTG at 20˚C, Lane 3: 400 µM IPTG at 20˚C, Lane 4: 600 µM IPTG at 20˚C, Lane 5: 0 µM IPTG at 37˚C (uninduced sample), Lane 6: 200 µM IPTG at 37˚C, Lane 7: 400 µM IPTG at 37˚C, Lane 8: 600 µM IPTG at 37˚C. (A) Soluble protein fractions. The arrow indicates to the possible overexpressed MBP band. (B) Insoluble protein fractions.

Consequently, it was decided to perform expression trials using only the pMALcHT-*msDPCK* plasmid to see if it could produce a soluble MsDPCK_MBP-His protein (~65 kDa). These expression trials were performed under the same conditions as the expression trials with the double plasmid containing *E. coli* cells. The SDS-PAGE analysis of these expression trials (Figure 4.17) showed that the MsDPCK_MBP-His protein was indeed overexpressed at 37˚C, albeit in the insoluble fractions. This can be seen by the visible overexpressed protein bands (indicated by the arrow) in the induced samples at 37˚C (Lane 6-8) of the insoluble protein fractions (Figure 4.17 B) and the lack of overexpressed protein bands in the soluble protein fractions (Figure 4.17 A).

Following guidelines for improving the solubility of MBP-fused proteins by Waugh [241], alternative expression trials were performed by, once again, using only the pMALcHT-*msDPCK* plasmid for expression of the MsDPCK_MBP-His protein. However, similar to the previous expression trials, the production of soluble protein was unsuccessful. This result was further confirmed by Dot-blot analysis.

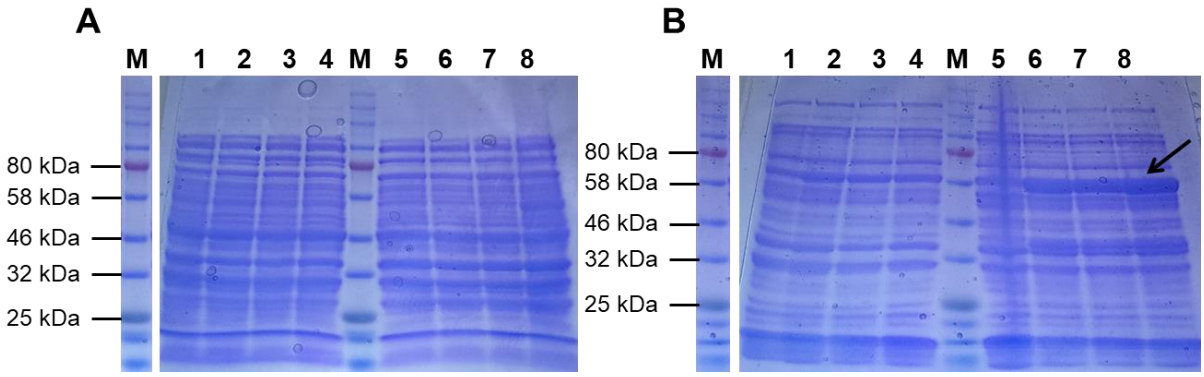


Figure 4.17 12% SDS-PAGE analysis of O/N expression trials at 20˚C and 37˚C using only the pMALcHT-*msDPCK* plasmid for expression of MsDPCK_MBP-His protein. M: Broad Range Protein Standard marker, Lane 1: 0 µM IPTG at 20˚C (uninduced sample), Lane 2: 200 µM IPTG at 20˚C, Lane 3: 400 µM IPTG at 20˚C, Lane 4: 600 µM IPTG at 20˚C, Lane 5: 0 µM IPTG at 37˚C (uninduced sample), Lane 6: 200 µM IPTG at 37˚C, Lane 7: 400 µM IPTG at 37˚C, Lane 8: 600 µM IPTG at 37˚C. (A) Soluble protein fractions. (B) Insoluble protein fractions. The arrow indicates to the possible overexpressed MsDPCK_MBP-His protein band.

4.4.4.4 Molecular chaperone-assisted expression

The SDM_pET28a(+)-*msDPCK*n plasmid, in combination with the pG-KJE8 plasmid, was used for the co-expression of MsDPCK_N-His and five molecular chaperones, namely GroEL (±60 kDa), GroES (±10 kDa), DnaK (±70 kDa), DnaJ (±40 kDa) and GrpE (±22 kDa). All of the conditions tested with the expression trials (as mentioned in Section 4.3.5.1) using the pG-KJE8 plasmid for molecular chaperone co-expression, failed to produce a soluble MsDPCK_N-His protein. Dot-blot analyses of the expression trials provided additional confirmation of this recurring result.

The SDS-PAGE analysis of the soluble and insoluble protein fractions obtained after the 2 h expression trials at 37˚C is shown (Figure 4.18) as an example of the failed solubility attempts via co-expression with molecular chaperone. The absent target protein band in the soluble protein fractions, specifically the double plasmid induced sample (Lane 5, Figure 4.18 A), indicates that no soluble overexpression occurred. Conversely, in the insoluble protein fractions (Figure 4.18 B) overexpression of the target protein did occur, which could be observed by the present protein band (indicated by the black arrow) in the double plasmid induced sample (Lane 5). The positive expression control was also successful as a protein band at the expected size (indicated by the orange arrow) can be observed in the insoluble fraction (Lane 2, Figure 4.18 B). Comparison of the positive control protein band (Lane 2) and the double plasmid induced sample protein band (Lane 5) suggest that these two proteins are the same proteins, i.e. MsDPCK_N-His.

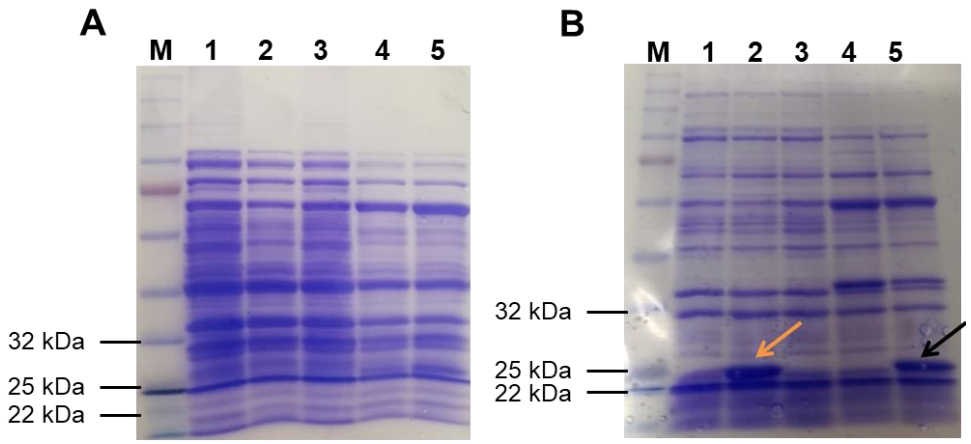


Figure 4.18 12% SDS-PAGE analysis of 2 h expression trials at 37˚C using the double plasmid containing *E. coli* cells for expression of MsDPCK_N-His and molecular chaperones. The expression of MsDPCK_N-His with *E. coli* cells containing only the SDM_pET28a(+)-*msDPCK*n plasmid, was used as a positive expression control. M: Broad Range Protein Standard marker, Lane 1: 0 µM IPTG for MsDPCK_N-His (uninduced positive control), Lane 2: 400 µM IPTG for MsDPCK_N-His (induced positive control), Lane 3: 0 µM IPTG, 0 mg/ml arabinose and 0 ng/ml tetracycline for chaperones and MsDPCK_N-His (uninduced sample of both plasmids), Lane 4: 0 µM IPTG, 0.5 mg/ml arabinose and 5 ng/ml tetracycline for chaperones and MsDPCK_N-His (uninduced sample of SDM_pET28a(+)-*msDPCK*n plasmid), Lane 5: 400 µM IPTG, 0.5 mg/ml arabinose and 5 ng/ml tetracycline for chaperones and MsDPCK_N-His. (A) Soluble protein fractions. (B) Insoluble protein fractions. The orange arrow indicates to the overexpressed MsDPCK_N-His protein band of the positive control. The black arrow indicates to an overexpressed protein band, which could possibly be MsDPCK_N-His.

4.4.5 Large-scale expression with sarkosyl detergent treatment

After the numerous failed attempts at soluble expression of the MsDPCK protein, it was decided to perform sarkosyl detergent treatment during cell lysis, following expression of the MsDPCK_N-His protein. The SDS-PAGE analysis of the large-scale expression samples (Figure 4.19 A), including the supernatant of the lysate after sarkosyl treatment, revealed that the target protein was overexpressed and soluble. This can be seen by the visible protein band (indicated by the arrow) in the supernatant sample of the lysate subsequent to treatment with 2% (v/v) sarkosyl. The corresponding Western-blot analysis of these samples (Figure 4.19 B) exhibited a positive signal for the 6xHis-tag at the expected size (orange box) in the three samples taken after IPTG induction (Lane 2-4). This provided additional confirmation that the overexpressed MsDPCK_N-His protein (Lane 4) was present in the soluble fraction (supernatant of the cell lysate).

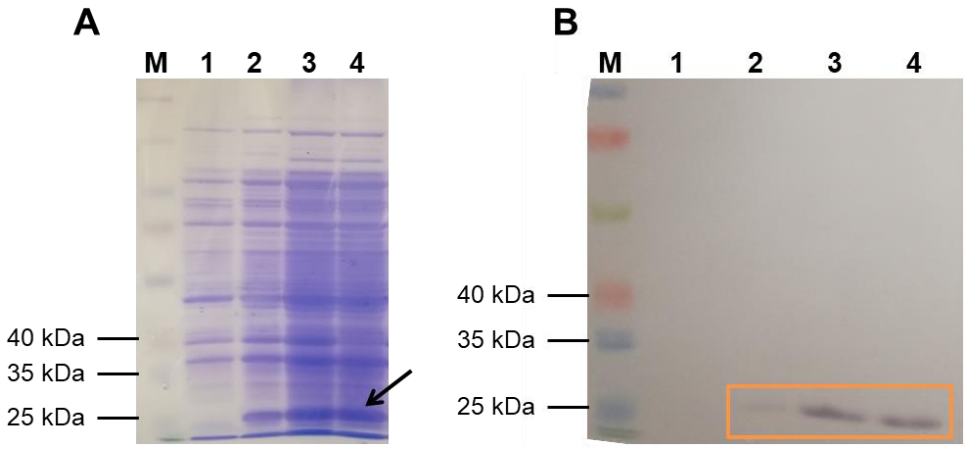


Figure 4.19 Analysis of large-scale expression followed by sarkosyl detergent treatment of the harvested cells. M: Spectra™ Multicolor Broad Range Protein marker, Lane 1: Sample prior to induction with IPTG, Lane 2: Sample after induction but prior to cell harvest, Lane 3: Sample of cell lysate (debris included), Lane 4: Sample of lysate supernatant (after sarkosyl treatment). (A) 12% SDS-PAGE analysis of the expression samples. The arrow indicates to the soluble overexpressed MsDPCK_N-His. (B) Corresponding Western-blot analysis of the expression samples. The orange box indicates the positive signal for the 6xHis-tag.

4.4.6 Purification of Ms02 DPCK

Following the large-scale expression of MsDPCK_N-His with subsequent sarkosyl treatment, the supernatant of the cell lysate was used for purification of the target protein using IMAC. The SDS-PAGE analysis of the purified samples (Figure 4.20 A) showed that the MsDPCK_N-His protein was not entirely purified since more than one protein band could be observed (Lane 1-3). Subsequently, a Western-blot analysis of the same samples (Figure 4.20 B) revealed that the contaminating protein does not contain a His-tag as it failed to produce a signal using the HisProbe-HRP antibody. A positive signal was, however, observed at more or less the expected size (indicated by the arrow) that corresponds with the protein band observed in the SDS-PAGE analysis, which is believed to be the MsDPCK_N-His protein. Moreover, the pantothenate kinase (PanK) of *Staphylococcus aureus* containing a 6xHis-tag (previously purified and concentrated by a co-worker) was included as positive control. A positive signal was also observed for the positive control (Lane 4), which further support the reliability of the Western-blot analysis.

In an attempt to obtain a pure isolated MsDPCK_N-His protein, the purification wash step was optimised. This was unsuccessful as the SDS-PAGE analyses of all the optimisation attempts showed the presence of the contaminating protein. A high imidazole concentration resulted in less target protein with the contaminating protein still present, whilst a very low concentration resulted in additional impurities, other than the unwanted contaminating protein. An imidazole of 20 mM was, thus, selected as the optimal imidazole concentration for the wash step as it provided the most target protein with the least amount of impurities.

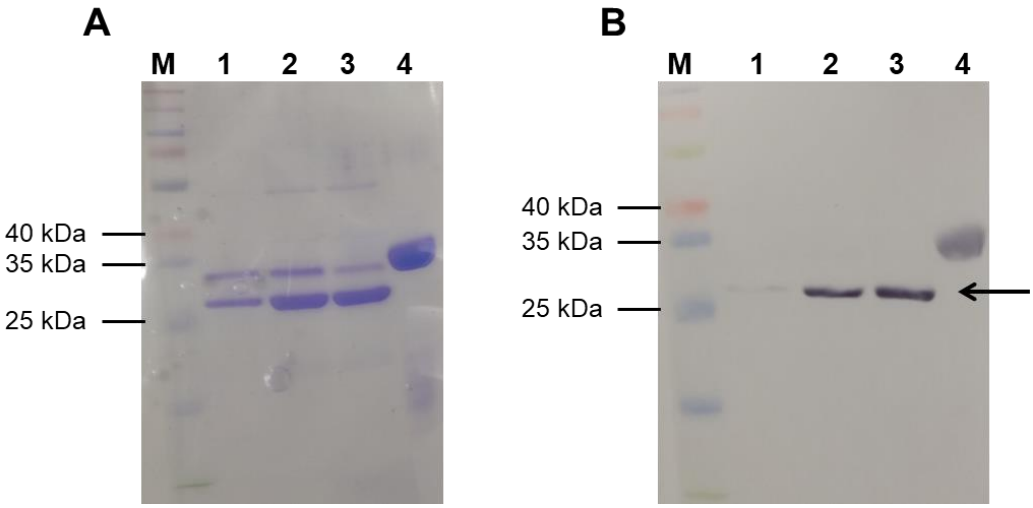


Figure 4.20 Analysis of purified samples subsequent to large-scale expression of MsDPCK_N-His and sarkosyl treatment. The purified and concentrated *S. aureus* PanK (with 6xHis-tag) was included as positive control. M: Spectra™ Multicolor Broad Range Protein marker, Lane 1-3: Respective purified samples of a single large-scale expression, Lane 4: Positive control. (A) 12% SDS-PAGE analysis of the purified samples. (B) Corresponding Western-blot analysis of the purified samples. The arrow indicates the positive signal of the purified MsDPCK_N-His protein.

A non-reducing SDS-PAGE analysis (Figure 4.21 B) was performed to determine if the recurring presence of the contaminating protein was because of association with the MsDPCK_N-His protein since obtaining a pure MsDPCK_N-His sample without the presence of the contaminating protein seemed to be unattainable using the applied purification method. A visible difference could be seen when a standard reducing SDS-PAGE analysis (Figure 4.21 A) and a non-reducing SDS-PAGE analysis (Figure 4.21 B) of the same purified samples were compared. In the non-reducing SDS-PAGE analysis, a distinct protein band with a size of ±46 kDa (orange box, Figure 4.21 B), along with the other two bands, could be observed in all of the purified samples. However, due to the small size difference of the MsDPCK_N-His protein and the contaminating protein, it is unclear whether this protein band is due to the association of the contaminating protein with the MsDPCK_N-His protein or whether it is due to dimer formation of the MsDPCK_N-His protein.

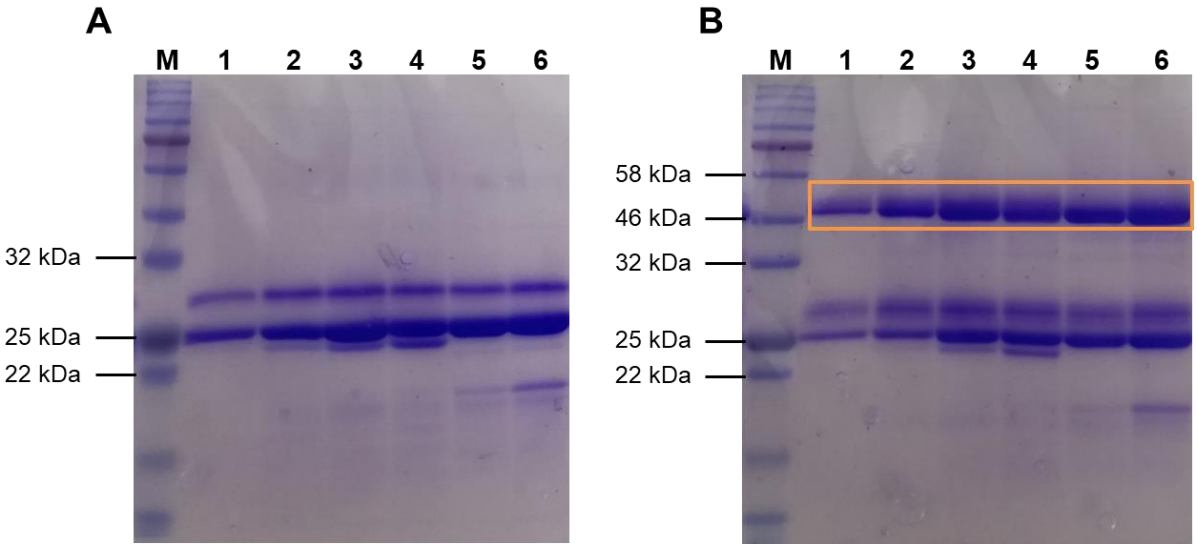


Figure 4.21 12% SDS-PAGE analysis of purified samples of three separate expression batches. M: Broad Range Protein Standard marker, Lane 1-2: Purified samples of expression batch 1, Lane 3-4: Purified samples of expression batch 2, Lane 5-6: Purified samples of expression batch 3. (A) Standard SDS-PAGE in which samples were treated with reducing treatment buffer. (B) Non-reducing SDS-PAGE in which samples were treated with non-reducing treatment buffer. The orange box indicates the protein band of the possible MsDPCK_N-His dimer or associated protein combination.

4.4.7 Preliminary test for activity of Ms02 DPCK

Even though the MsDPCK protein could not be isolated without the presence of an unknown protein contaminant, it was decided to use the isolate for determining possible enzymatic activity of the DPCK protein. The six purified samples that were obtained from three separate expression batches (two samples per batch as indicated by Figure 4.21), were used to determine if the expressed MsDPCK enzyme was active.

The subsequent HPLC analyses of Sample 5 (sample of Lane 5 in Figure 4.21 A) and the accompanying reaction controls are shown in Figure 4.22. The retention time of CoA (indicated by the red dotted line) was almost 7 min, whilst the retention time of DePCoA (indicated by the blue dotted line) was 11 min. This provided a clear distinction between the substrate (DePCoA) and product (CoA) and could, therefore, be used as an indication of enzymatic activity. An active enzyme will convert substrate into product, thus, resulting in a peak at the CoA retention time. Conversely, an inactive enzyme will not convert substrate into product, resulting in a peak at the DePCoA retention time only. In light of this, the HPLC analysis of Sample 5 (Figure 4.22 D) suggests that the purified MsDPCK is inactive since only a peak at the DePCoA retention time could be observed. This was the case for all six samples. The *E. coli* DPCK (positive control) did, however, exhibit activity as the HPLC analysis thereof (Figure 4.22 C) displayed a peak at the CoA retention time.

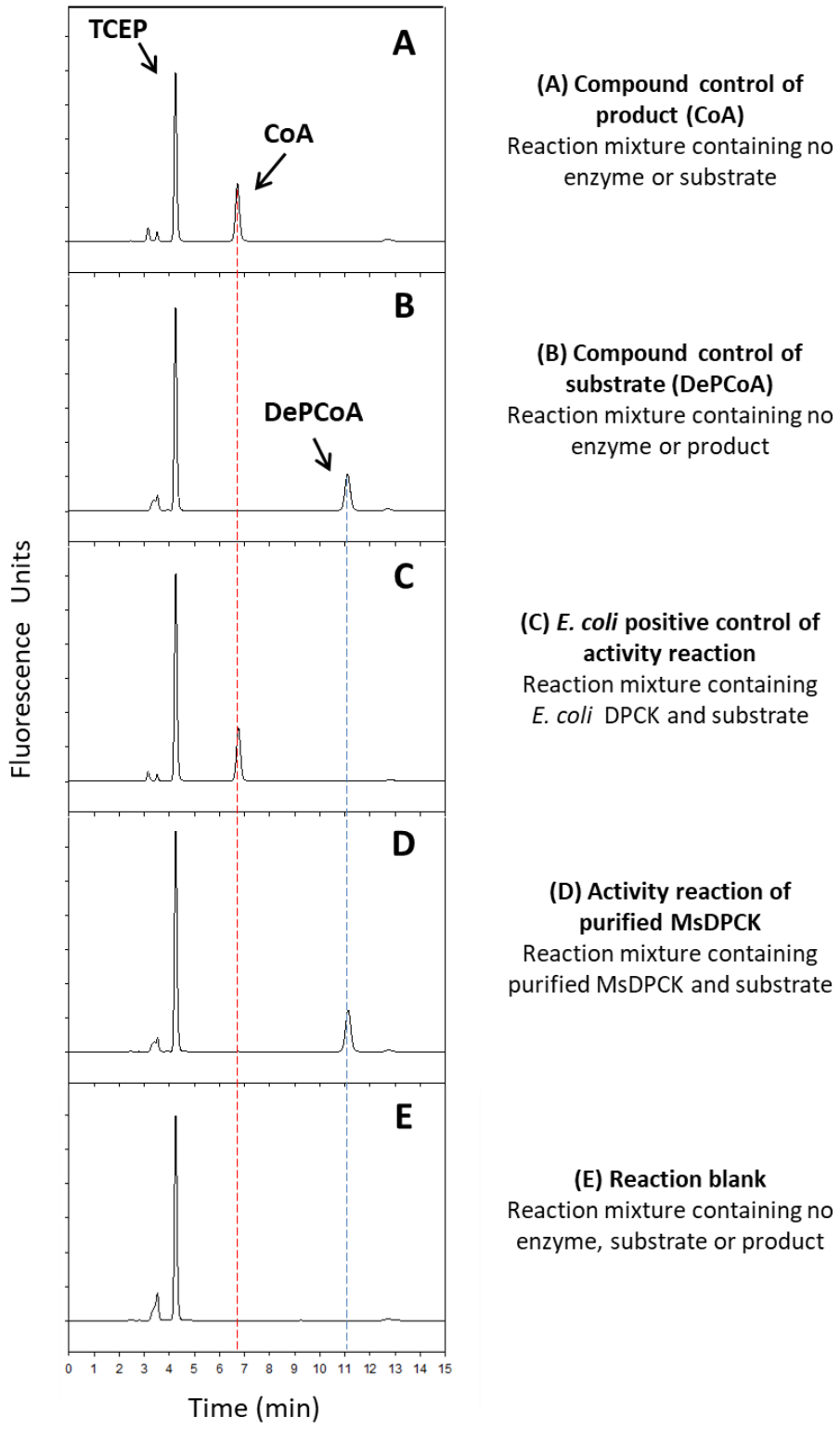


Figure 4.22 HPLC analyses of the enzymatic reactions of the purified MsDPCK samples. The retention time of CoA is indicated by the red dotted line and the retention time of DePCoA is indicated by the blue dotted line. The reaction mixture variables are shown on the right of the relevant analysis. (A) Compound control of product (CoA). (B) Compound control of substrate (DePCoA). (C) *E. coli* positive control of activity reaction. (D) Activity reaction of purified MsDPCK (Sample 5). (E) Reaction blank.

4.5 Discussion

If the DPCK enzyme of Ms02 was to be considered as a suitable target for drug or vaccine development, it is important to have the necessary knowledge concerning the functionality and kinetic parameters of this enzyme. Hence, in order to allow the future characterisation of the Ms02 DPCK enzyme, this protein needs to be produced and isolated. Production of the protein can be accomplished by recombinant expression along with a fusion tag which will allow subsequent isolation.

In this study, the *msDPCK* gene was successfully amplified from Ms02 genomic DNA and cloned into a pET28a(+) expression vector for expression with an N-terminal 6xHis-tag. Because of the difference in the genetic codes that are utilised by mycoplasmas compared that of other prokaryotes [22, 33], SDM was required to ensure that no premature stop codons are present in the *msDPCK* gene insert. Consequently, the two codons in question were successfully mutated and the resulting SDM_pET28a(+)-*msDPCK*n plasmid transformed into an *E. coli* expression strain.

After the initial large-scale expression and purification of the MsDPCK_N-His protein, it was discovered that this protein was insolubly expressed. According to literature, the solubility of recombinant proteins can be increased by optimising expression conditions through varying the growth medium, induction temperature, induction time and inducer concentration [254, 255]. The subsequent expression trials, however, still resulted in the insoluble expression of the target protein. This prompted an investigation into alternative expression approaches that might result in the soluble expression of the MsDPCK protein.

Since the position of a His-tag on its fused protein partner might have adverse effects on the protein structure and properties [233], one of the approaches was to shift the 6xHis-tag from its existing position at the N-terminus to the C-terminus. Accordingly, the SDM_*msDPCK* gene was amplified from the SDM_pET28a(+)-*msDPCK*n and sub-cloned into a pET28a(+) expression vector for expression with a C-terminal 6xHis-tag. Regardless of the correctly inserted gene, the overexpression of the target protein could not be achieved in the expression trials using this vector construct. This lack of overexpression was not due to faulty expression conditions since the positive control displayed overexpression. The addition of a C-terminal His-tag has previously been shown to have an unexpected influence on the properties of the fused protein [256]. Thus, this lack of overexpression observed subsequent to the repositioning of the His-tag provides additional evidence that the position of a His-tag can have major influences on the protein properties. In this case, the terminal shift of the 6xHis-tag seemed to be detrimental to the expression of the target protein.

The fusion of certain affinity tags with proteins is well-known for enhancing the solubility of the attached protein [234, 235]. Hence, as an attempt to produce a soluble MsDPCK protein, it was fused with a modified MBP-6xHis-tag. The specific modified fusion tag used in this study allowed for the cleaving of the MBP domain via TEV protease digestion, producing a 6xHis-tagged protein. Overexpression of the cleaved target protein could, however, not be achieved. However, the SDS-PAGE analysis of the soluble fractions did exhibit a protein band, in the induced samples, at a size corresponding to that of the MBP domain (Figure 4.16 A). If this observed protein band was indeed the MBP domain, it would suggest that the MsDPCK_MBP-His protein was successfully expressed and cleaved; yet there was no visible band relating to the cleaved target protein in the soluble protein fractions.

In order to determine if the MsDPCK_MBP-His protein would express soluble without immediate cleavage of the MBP domain, expression trials using only the pMALcHT-*msDPCK* plasmid were performed. The MsDPCK_MBP-His protein was shown to be overexpressed; however, it was observed in the insoluble fraction. The same result was observed after following the guidelines proposed by Waugh [241] to improve solubility of MBP fused proteins. Therefore, regardless of the MBP domain being soluble on its own, it could not improve the solubility of the MsDPCK protein.

The use of molecular chaperones to improve the solubility of recombinant proteins have been reported in several studies [245–249]. Consequently, the pG-KJE8 plasmid was transformed into Invitrogen™ *E. coli* BL21 Star™ (DE3) cells that already contain the SDM_pET28a(+)-*msDPCK*n plasmid, followed by the co-expression of the DnaK, DnaJ, GrpE, GroES, and GroEL molecular chaperones with the MsDPCK_N-His protein. However, the various expression trials performed using these plasmids, could not produce a soluble target protein. This is most probably due to misfolding of the protein, resulting in aggregation, which is (according to Thomas *et al.* [245]) a possible outcome in the proposed model for chaperone-assisted protein folding in *E. coli* (Figure 4.3).

The recurring failed attempts at soluble expression of MsDPCK lead to the alternative approach of sarkosyl treatment, since sarkosyl has reportedly been used to solubilise recombinant proteins from inclusion bodies [250, 251]. The MsDPCK_N-His protein was selected for this trial because it displayed the most stable expression of all the tested expression alternatives and would still allow purification by IMAC due to the fused 6xHis-tag. The sarkosyl treatment was successful in solubilising the MsDPCK_N-His protein, which could subsequently be purified.

The MsDPCK_N-His protein could not be isolated in a pure form with the method used in this study. The presence of an unknown contaminating protein was repeatedly observed following multiple purification attempts, but western-blot analysis confirmed that this protein does not contain a His-tag. Hence, it was thought that this contaminating protein might associate closely with the target protein, thus eluting at the same time. A non-reducing SDS-PAGE analysis was performed in order to test this hypothesis. Although a protein band was observed at ±46 kDa (orange box, Figure 4.21 B), the results were inconclusive since there is a small size difference between these two proteins and the two bands seen in the reducing SDS-PAGE analysis (Figure 4.21 A) could still be observed in the non-reducing SDS-PAGE analysis. Hence, the two protein bands did not become one, which would indicate protein association; nor did the smaller target protein band disappear, which would indicate dimerisation.

However, the comparison of the SDS-PAGE analysis of the initial MsDPCK_N-His expression and purification (Figure 4.13 A) to the sarkosyl-solubilised MsDPCK_N-His purifications (Figure 4.21 A) revealed that the unknown contaminating protein was already present in the soluble protein fraction of initial expression, which can be seen in the purified samples (Lane 3-4, Figure 4.13 A). Therefore, since the MsDPCK_N-His protein was only present in the insoluble protein fraction of the initial expression, it suggests that the contaminating protein cannot be an association protein because only the soluble protein fraction could be purified. This implies that this contaminating protein is merely a His-rich protein that binds equally or possibly better to the column than the MsDPCK_N-His protein.

Moreover, the observed protein band at ±46 kDa (orange box) in the non-reducing SDS-PAGE analysis (Figure 4.21 B) suggests that the MsDPCK_N-His protein forms dimers in solution. Thus far, no literature has reported the dimerisation of DPCK. The DPCK of *E. coli* (usually a monomer [156]) has, however, been reported to crystallise as a trimer, albeit in the presence of sulphate [157]. On the other hand, this dimer formation might be due to the addition of the 6xHis-tag, which has been reported to cause unexpected dimerisation of its (normally monomeric) protein fusion partner [257].

An HPLC analysis method, developed by Goosen *et al*. [231], was used to test for DPCK enzymatic activity of the purified protein samples (Figure 4.22). This method was successful in determining whether an enzyme is active or not as the conversion of substrate to product could be detected (as seen in the positive control). None of the six tested samples contained an active MsDPCK protein. Since sarkosyl is a detergent, using it to solubilise the MsDPCK_N-His protein might have an influence on the activity of the protein. However, according to Massiah *et al*. [252], the use of sarkosyl treatment to solubilise proteins does

not affect the functional properties of the proteins and can be used in biological assays and structure analyses.

There are many reasons for the lack of activity observed in purified protein samples, such as: the protein is not folded properly in the *E. coli* expression host; the protein needs to undergo post-translational modifications to become active in Ms02, which is missing in the *E. coli* expression host; the 6xHis-tag has a detrimental effect on the functional properties of the protein [232, 233]; or the protein has been inactivated as a result of ineffective handling of the purified protein [258]. The latter can be due to a change in solution conditions, a change in physical conditions (e.g. freezing and thawing), exposure to heavy metals, oxygen, and degradative enzymes, or inadequate storage conditions, such as the sample being too diluted or insufficient glycerol added. Thus, any one of these reasons could be responsible for the observed inactivity of the MsDPCK samples.

Due to the impure isolates and lack of DPCK enzymatic activity, the characterisation of the MsDPCK enzyme could not be performed during this study. However, the *msDPCK* gene was successfully amplified, cloned and mutated with SDM. Moreover, an array of expression trials was performed in an attempt to solubilise the MsDPCK protein. Therefore, the mutated gene, along with the information regarding the solubility of the expressed protein, serves as basis for future studies, which could subsequently allow for characterisation of the MsDPCK enzyme.

Chapter 5 – Concluding remarks and future perspectives

With the purpose of identifying enzymes within the coenzyme A (CoA) biosynthesis pathway that would be suitable as potential drug or vaccine targets against mycoplasma infections in ostriches, the first aim of this study was to determine the presence or absence of enzyme-encoding genes in the CoA biosynthesis pathway of *Mycoplasma* species. Using a bioinformatics approach, 62 species were investigated of which eight species (13%) were found to contain none of the pathway enzyme-encoding genes. The remaining species all possessed at least one of the CoA biosynthetic pathway enzyme-encoding genes. Twelve enzyme-encoding gene homologues (ten hypothetical and two putative) were identified and their predicted identities were confirmed by determining the conserved and functional motifs and domains.

Many of the genomes of the *Mycoplasma* species that were investigated were only available at contiguous sequence level. This might influence the number of CoA biosynthetic pathway enzyme-encoding genes that were identified, since there might be genomic information missing in these genomes. Nevertheless, the most frequent enzyme-encoding gene present in the investigated *Mycoplasma* species was that of DPCK, which suggests that DPCK has a more prominent role in the biosynthesis of CoA amongst these species. Therefore, this enzyme should be the starting point of an investigation into anti-mycoplasmal agents of a particular species.

Furthermore, there seemed to be no correlation between the number of identified CoA biosynthesis pathway enzyme-encoding genes in a species and the phylogeny of the respective proteins. There was also no correlation with the 16S rRNA phylogenetic groupings, which is currently recommended for species classification. This implies that an exhaustive investigation into the CoA biosynthesis pathway of the mycoplasma, for which an anti-mycoplasmal agent is to be developed, should be performed to gain a better understanding of its CoA biosynthesis requirements.

In future studies, the number of investigated species can be increased as this would provide an even larger variety of species, which could allow for a better understanding of the distribution of the enzyme-encoding genes associated with CoA biosynthesis. Since the depository of species genomes is rapidly increasing, along with the quality of genome sequencing and annotation, a future repeat of this study could also offer some answers to the genome related questions. Additionally, because the type III PanK enzyme (present in some of the *Mycoplasma* species) displayed variation from the norm in its active cite residues, the characterisation of this enzyme could provide some insight into the role of this

enzyme in these organisms. The same can be said for the unknown bifunctional HAD-like protein/dephospho-CoA kinase (HAD-DPCK) protein found in some of the species.

Taking in to consideration the importance of the DPCK enzyme as highlighted by the first aim, the second aim of this study was to recombinantly express and isolate the DPCK enzyme of the ostrich infecting *Mycoplasma* sp. Ms02 (Ms02). The amplification and cloning of the *msDPCK* gene was successful, along with its site-directed mutagenesis. The soluble expression of Ms02 DPCK protein (MsDPCK) proved to be challenging, but treatment with sarkosyl was able to solubilise the N-terminal 6xHis-tagged MsDPCK protein. A pure fraction of this protein could not be isolated due to the simultaneous elution of, what is thought to be, a His-rich protein. Irrespective of the presence of this contaminating protein, the purified samples were used for the preliminary testing of DPCK enzyme activity. However, none of the tested samples exhibited DPCK activity.

The addition of a 6xHis-tag to the target protein seemed to influence the ability to express a soluble protein. As a suggestion for future studies, the MsDPCK protein could be fused with a glutathione-*S*-transferase (GST) affinity tag allowing purification without the presence of the His-rich contaminating protein. Should solubility still be a problem, the GST-tagged protein may be solubilised using the sarkosyl treatment method as described by Massiah *et al*. [252]. However, if the inactivity of the MsDPCK enzyme is in fact due to the presence of sarkosyl in the samples, there are many other alternative expression approaches that could be investigated that might result in a soluble protein, such as using a different expression vector system or using a different expression host entirely. Moreover, the MsDPCK protein could also be expressed without a fused affinity tag. This will also provide additional information about the native structure of the MsDPCK protein (if soluble), which could provide answers concerning the possible formation of dimers in solution.

References

1. Marenda, M. S. (2014) Genomic mosaics. in *Mollicutes: Molecular Biology and Pathogenesis* (Browning, G. F., and Citti, C. eds), pp. 15–54, Caister Academic Press, Norfolk, UK

2. Razin, S., Yogev, D., and Naot, Y. (1998) Molecular biology and pathogenicity of mycoplasmas. *Microbiol. Mol. Biol. Rev.* **62**, 1094–156

3. Rivera, A., and Cedillo, L. (2015) Mycoplasmas and Nucleases. *J. Pure. Appl. Microbio.* **9**, 2115–25

4. Botes, A., Pretorius, B., Olivier, A. J., Burger, W. P., and Bellstedt, D. U. (2008) Vaccine strategies against ostrich mycoplasmas. *Aust. J. Exp. Agric. (Animal Prod. Sci. Spec. Issue Ratite Sci. Ind. Conserv.* **48**, ix

5. Botes, A., Peyrot, B. M., Olivier, A. J., Burger, W. P., and Bellstedt, D. U. (2005) Identification of three novel mycoplasma species from ostriches in South Africa. *Vet. Microbiol.* **111**, 159–69

6. Moolman, W. J., de Villiers, M., and Strauss, E. (2014) Recent advances in targeting coenzyme A biosynthesis and utilization for antimicrobial drug development. *Biochem. Soc. Trans.* **42**, 1080–6

7. Strauss, E. (2010) Coenzyme A Biosynthesis and Enzymology. in *Comprehensive Natural Products II: Volume 7: Cofactors* (Liu, H.-W., and Mander, L. eds), pp. 351–410, Elsevier Ltd., Oxford, UK

8. Spry, C., Kirk, K., and Saliba, K. J. (2008) Coenzyme A biosynthesis: an antimicrobial drug target. *FEMS Microbiol. Rev.* **32**, 56–106

9. Ye, Y., Osterman, A., Overbeek, R., and Godzik, A. (2005) Automatic detection of subsystem/pathway variants in genome analysis. *Bioinformatics*. **21**, i478-86

10. Hayflick, L., and Chanock, R. M. (1965) Mycoplasma Species of Man. *Bacteriol. Rev.* **29**, 185–221

11. Razin, S. (2010) Time-line of significant contributions to mycoplasmology. *Biologicals*. **38**, 191–2

12. Sabin, A. B. (1941) The Filtrable Microorganisms of the Pleuropneumonia Group. *Bacteriol. Rev.* **5**, 1–67

13. Razin, S., and Hayflick, L. (2010) Highlights of mycoplasma research—an historical perspective. *Biologicals*. **38**, 183–90

14. Edward, D. G., and Freundt, E. A. (1956) The classification and nomenclature of organisms of the pleuropneumonia group. *Microbiology*. **14**, 197–207

15. McGee, Z. A., Rogul, M., and Wittler, R. G. (1967) Molecular genetic studies of relationships among Mycoplasma, L-forms and Bacteria. *Ann. N. Y. Acad. Sci.* **143**, 21–30

16. Somerson, N. L., Reich, P. R., Chanock, R. M., and Weissman, S. M. (1967) Genetic differentiation by nucleic acid homology. III. Relationships among Mycoplasma, L-forms, and Bacteria. *Ann. N. Y. Acad. Sci.* **143**, 9–20

17. Morowitz, H. J., and Wallace, D. C. (1973) Genome size and life cycle of the mycoplasma. *Ann. N. Y. Acad. Sci.* **225**, 62–73

18. Neimark, H. C. (1986) Origin and evolution of wall-less prokaryotes. in *The bacterial L-forms* (Madoff, S. ed), pp. 21–42, Marcel Dekkar Inc., New York, USA

19. Woese, C. R., Maniloff, J., and Zablen, L. B. (1980) Phylogenetic analysis of the mycoplasmas. *Proc. Natl. Acad. Sci. U. S. A.* **77**, 494–8

20. Fraser, C. M., Gocayne, J. D., White, O., Adams, M. D., Clayton, R. A., Fleischmann, R. D., Bult, C. J., Kerlavage, A. R., Sutton, G., Kelley, J. M., Fritchman, J. L., Weidman, J. F., Small, K. V., Sandusky, M., Furhmann, J., Nguyen, D., Utterback, T. R., Saudek, D. M., Phillips, C. A., Merrick, J. M., Tomb, J.-F., Dougherty, B. A., Bott, K. F., Hu, P.-C., Lucier, T. S., Peterson, S. N., Smith, H. O., Hutchison III, C. A., and Venter, J. C. (1995) The minimal gene complement of Mycoplasma genitalium. *Science.* **270**, 397–403

21. Razin, S. (2006) The genus Mycoplasma and related genera (class Mollicutes). in *The Prokaryotes: Volume 4: Bacteria: Firmicutes, Cyanobacteria*, 3rd Ed. (Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K., and Stackebrandt, E. eds), pp. 836–904, Springer, New York, USA

22. Le Roux, M. C., and Hoosen, A. A. (2010) Mycoplasma genitalium: a brief review. *South. African J. Epidemiol. Infect.* **25**, 7–10

23. Razin, S. (2005) Mycoplasma. in *Topley and Wilson's microbiology and microbial infections: Volume 2: Bacteriology*, 10th Ed. (Borriello, S. P., Murray, P. R., and Funke, G. eds), pp. 1957–2005, Hodder Arnold and ASM Press, London, UK

24. Bradbury, J. M. (2005) Poultry mycoplasmas: sophisticated pathogens in simple guise. *Br. Poult. Sci.* **46**, 125–36

25. Johansson, K.-E., and Pettersson, B. (2002) Taxonomy of mollicutes. in *Molecular Biology and Pathogenicity of Mycoplasmas* (Razin, S., and Herrmann, R. eds), pp. 1–29, Kluwer Academic Publishers, New York, USA

26. Brown, D. R. (2010) Phylum XVI. Tenericutes Murray 1984a, 356VP (Effective publication: Murray 1984b, 33.). in *Bergey's Manual® of Systematic Bacteriology: Volume 4*, 2nd Ed. (Whitman, W. B., and Parte, A. C. eds), pp. 567–723, Springer, New York, USA

27. Kamla, V., Henrich, B., and Hadding, U. (1996) Phylogeny based on elongation factor Tu reflects the phenotypic features of mycoplasmas better than that based on 16S rRNA. *Gene*. **171**, 83–7

28. Maniloff, J. (2002) Phylogeny and evolution. in *Molecular Biology and Pathogenicity of Mycoplasmas* (Razin, S., and Herrmann, R. eds), pp. 31–43, Kluwer Academic Publishers, New York, USA

29. Weisburg, W. G., Tully, J. G., Rose, D. L., Petzel, J. P., Oyaizu, H., Yang, D., Mandelco, L., Sechrest, J., Lawrence, T. G., and Van Etten, J. (1989) A phylogenetic analysis of the mycoplasmas: basis for their classification. *J. Bacteriol.* **171**, 6455–67

30. Johansson, K.-E., Heldtander, M. U. K., and Pettersson, B. (1998) Characterization of Mycoplasmas by PCR and Sequence Analysis with Universal 16S rDNA Primers. in *Mycoplasma Protocols* (Miles, R., and Nicholas, R. eds), pp. 145–65, Humana Press, Totowa, USA

31. Fadiel, A., Lithwick, S., and Naftolin, F. (2005) The influence of environmental adaptation on bacterial genome structure. *Lett. Appl. Microbiol.* **40**, 12–8

32. Waites, K. B., Xiao, L., Paralanov, V., Viscardi, R. M., and Glass, J. I. (2013) Mycoplasma and ureaplasma. in *Molecular Typing in Bacterial Infections* (de Filippis, I., and McKee, M. L. eds), pp. 229–81, Humana Press, New York, USA

33. Taylor-Robinson, D., and Jensen, J. S. (2011) Mycoplasma genitalium: from Chrysalis to multicolored butterfly. *Clin. Microbiol. Rev.* **24**, 498–514

34. Himmelreich, R., Plagens, H., Hilbert, H., Reiner, B., and Herrmann, R. (1997) Comparative analysis of the genomes of the bacteria Mycoplasma pneumoniae and Mycoplasma genitalium. *Nucleic Acids Res.* **25**, 701–12

35. Henrich, B., Hopfe, M., Kitzerow, A., and Hadding, U. (1999) The adherence-associated lipoprotein P100, encoded by an opp operon structure, functions as the oligopeptide-binding domain OppA of a putative oligopeptide transport system in Mycoplasma hominis. *J. Bacteriol.* **181**, 4873–8

36. Pollack, J. D. (2002) Central carbohydrate pathways: Metabolic flexibility and the extra role of some “housekeeping” enzymes. in *Molecular Biology and Pathogenicity of Mycoplasmas* (Razin, S., and Herrmann, R. eds), pp. 163–99, Kluwer Academic Publishers, New York, USA

37. Razin, S. (1978) The mycoplasmas. *Microbiol. Rev.* **42**, 414–70

38. Blanchard, A., and Bébéar, C. M. (2002) Mycoplasmas of humans. in *Molecular Biology and Pathogenicity of Mycoplasmas* (Razin, S., and Herrmann, R. eds), pp. 45–71, Kluwer Academic Publishers, New York, USA

39. Frey, J. (2002) Mycoplasmas of animals. in *Molecular Biology and Pathogenicity of Mycoplasmas* (Razin, S., and Herrmann, R. eds), pp. 73–90, Kluwer Academic Publishers, New York, USA

40. Brown, D. R., Farley, J. M., Zacher, L. A., Carlton, J. M., Clippinger, T. L., Tully, J. G., and Brown, M. B. (2001) Mycoplasma alligatoris sp. nov., from American alligators. *Int. J. Syst. Evol. Microbiol.* **51**, 419–24

41. Seemüller, E., Garnier, M., and Schneider, B. (2002) Mycoplasmas of plants and insects. in *Molecular Biology and Pathogenicity of Mycoplasmas* (Razin, S., and Herrmann, R. eds), pp. 91–115, Kluwer Academic Publishers, New York, USA

42. Tully, J. G. (1996) Mollicute-host interrelationships: Current concepts and diagnostic implications. *Mol. diagnostic Proced. Mycoplasmol.* **2**, 1–21

43. Hackett, K. J., and Clark, T. B. (1989) Ecology of spiroplasmas. *The mycoplasmas*. **5**, 113–200

44. Langer, S. (2009) *Proposal for molecular tools for the epidemiology of Contagious Bovine Pleuro Pneumonia and classification of unknown Mycoplasma sp. isolated from Struthio camelus.* MSc thesis, Uni. Vienna

45. Wium, M., Botes, A., and Bellstedt, D. U. (2015) The identification of oppA gene homologues as part of the oligopeptide transport system in mycoplasmas. *Gene*. **558**, 31–40

46. Huchzermeyer, F. W. (1994) *Ostrich diseases*, Onderstepoort Veterinary Institute

47. Kleven, S. H. (2008) Control of Avian Mycoplasma Infections in Commercial Poultry. *Avian Dis.* **52**, 367–74

48. Nascimento, E. R., Pereira, V. L. A., Nascimento, M. G. F., and Barreto, M. L. (2005) Avian mycoplasmosis update. *Rev. Bras. Ciência Avícola*. **7**, 1–9

49. de Wet, B. (2015) *Evaluation of DNA vaccines developed against Mycoplasma struthionis sp. nov. str. Ms01 in ostriches*. MSc thesis, Uni. Stellenbosch

50. Stanley, W. A., Hofacre, C. L., Speksnijder, G., Kleven, S. H., and Aggrey, S. E. (2001) Monitoring Mycoplasma gallisepticum and Mycoplasma synoviae infection in breeder chickens after treatment with enrofloxacin. *Avian Dis.* **45**, 534–39

51. Reinhardt, A. K., Gautier-Bouchardon, A. V, Gicquel-Bruneau, M., Kobisch, M., and Kempf, I. (2005) Persistence of Mycoplasma gallisepticum in chickens after treatment with enrofloxacin without development of resistance. *Vet. Microbiol.* **106**, 129–37

52. Umar, S., Munir, M. T., Ur-Rehman, Z., Subhan, S., Azam, T., and Shah, M. A. A. (2017) Mycoplasmosis in poultry: update on diagnosis and preventive measures. *Worlds. Poult. Sci. J.* **73**, 17–28

53. Leigh, S. A., Branton, S. L., Evans, J. D., and Collier, S. D. (2013) Impact of fowlpox-vectored Mycoplasma gallisepticum vaccine Vectormune FP MG on layer hen egg production and egg quality parameters. *Poult. Sci.* **92**, 3172–5

54. Ferguson-Noel, N., Cookson, K., Laibinis, V. A., and Kleven, S. H. (2012) The efficacy of three commercial Mycoplasma gallisepticum vaccines in laying hens. *Avian Dis.* **56**, 272–5

55. Whithear, K. G. (1996) Control of avian mycoplasmoses by vaccination. *Rev. Sci. Tech.* **15**, 1527–53

56. Pretorius, B. (2009) *The Ostrich Mycoplasma Ms01: The Identification Isolation And Modification Of The P100 Vaccine Candidate Gene And Immunity Elicited By Poultry Mycoplasma Vaccines*. MSc thesis, Uni. Stellenbosch

57. Steenmans, S. (2010) *The identification, contiguous sequence annotation, cloning and site-directed mutagenesis of the P100 vaccine candidate gene of the ostrich mycoplasma Ms02*. MSc thesis, Uni. Stellenbosch

58. Wium, M. (2015) *The development of a DNA vaccine against Mycoplasma nasistruthionis sp. nov. for use in ostriches*. PhD thesis, Uni. Stellenbosch

59. Strydom, M. (2013) *The ostrich mycoplasma Ms02: Partial genome assembly, bioinformatic analysis and the development of three DNA vaccines*. MSc thesis, Uni. Stellenbosch

60. Brandt, S. (2012) *Preliminary investigations into the use of DNA vaccines to elicit protective immune responses against the ostrich mycoplasma Ms01*. MSc thesis, Uni. Stellenbosch

61. Van Tonder, A. (2013) *Development of an oral vaccine against the ostrich-specific mycoplasma, Mycoplasma struthionis*. MSc thesis, Uni. Stellenbosch

62. Jonker, H. I. (2017) *Evaluation of DNA vaccines against Mycoplasma nasistruthionis sp. nov. str. Ms03 infections in ostriches and the production of IgA heavy chain proteins*. MSc thesis, Uni. Stellenbosch

63. Butt, A. M., Tahir, S., Nasrullah, I., Idrees, M., Lu, J., and Tong, Y. (2012) Mycoplasma genitalium: a comparative genomics study of metabolic pathways for the identification of drug and vaccine targets. *Infect. Genet. Evol.* **12**, 53–62

64. Balish, M. F., and Distelhorst, S. L. (2016) Potential Molecular Targets for Narrow-Spectrum Agents to Combat Mycoplasma pneumoniae Infection and Disease. *Front. Microbiol.* **7**, 205

65. Lipmann, F., Kaplan, N. O., and Novelli, G. D. (1947) Chemistry and distribution of the coenzyme for acetylation (coenzyme A). *Fed. Proc.* **6**, 272

66. Leonardi, R., Zhang, Y.-M., Rock, C. O., and Jackowski, S. (2005) Coenzyme A: Back in action. *Prog. Lipid Res.* **44**, 125–53

67. Choudhary, C., Weinert, B., Nishida, Y., Verdin, E., and Mann, M. (2014) The growing landscape of lysine acetylation links metabolism and cell signalling. *Nat. Rev. Mol. Cell Biol.* **15**, 536–50

68. Pietrocola, F., Galluzzi, L., Pedro, J. B.-S., and Madeo, F. (2015) Acetyl Coenzyme A: A Central Metabolite and Second Messenger. *Cell Metab.* **21**, 805–21

69. Mishra, P. K., and Drueckhammer, D. G. (2000) Coenzyme A Analogues and Derivatives: Synthesis and Applications as Mechanistic Probes of Coenzyme A Ester-Utilizing Enzymes. *Chem. Rev.* **100**, 3283–310

70. Yang, W., and Drueckhammer, D. G. (2000) Computational Studies of the Aminolysis of Oxoesters and Thioesters in Aqueous Solution. *Org. Lett.* **2**, 4133–6

71. Yang, W., and Drueckhammer, D. G. (2001) Understanding the Relative Acyl-Transfer Reactivity of Oxoesters and Thioesters: Computational Analysis of Transition State Delocalization Effects. *J. Am. Chem. Soc.* **123**, 11004–9

72. Brown, G. M., and Reynolds, J. J. (1963) Biogenesis of the Water-Soluble Vitamins. *Annu. Rev. Biochem.* **32**, 419–62

73. Begley, T. P., Kinsland, C., and Strauss, E. (2001) The biosynthesis of coenzyme A in bacteria. *Vitam. Horm.* **61**, 157–71

74. Abiko, Y. (1975) Metabolism of coenzyme A. in *Metabolic Pathways: Volume VII: Metabolism of Sulfur Compounds*, 3rd Ed. (Greenberg, D. M. ed), pp. 1–25, Academic Press, New York, USA

75. de Villiers, M., Barnard, L., Koekemoer, L., Snoep, J. L., and Strauss, E. (2014) Variation in pantothenate kinase type determines the pantothenamide mode of action and impacts on coenzyme A salvage biosynthesis. *FEBS J.* **281**, 4731–53

76. Jackowski, S., and Rock, C. O. (1981) Regulation of coenzyme A biosynthesis. *J. Bacteriol.* **148**, 926–32

77. Maas, W. (1952) Pantothenate studies. III. Description of the extracted pantothenate-synthesizing enzyme of Escherichia coli. *J. Biol. Chem.* **198**, 23–32

78. Maas, W., and Novelli, G. (1953) Synthesis of pantothenic acid by depyrophosphorylation of adenosine triphosphate. *Arch. Biochem. Biophys.* **43**, 236–8

79. Maas, W. K., and Vogel, H. J. (1953) α-Ketoisovaleric acid, a precursor of pantothenic acid in Escherichia coli. *J. Bacteriol.* **65**, 388–93

80. Webb, M. E., Smith, A. G., and Abell, C. (2004) Biosynthesis of pantothenate. *Nat. Prod. Rep.* **21**, 695–721

81. Velaparthi, S., Brunsteiner, M., Uddin, R., Wan, B., Franzblau, S. G., and Petukhov, P. A. (2008) 5-tert-Butyl-N-pyrazol-4-yl-4,5,6,7-tetrahydrobenzo[d]isoxazole-3-carboxamide Derivatives as Novel Potent Inhibitors of Mycobacterium tuberculosis Pantothenate Synthetase: Initiating a Quest for New Antitubercular Drugs. *J. Med. Chem.* **51**, 1999–2002

82. Mageed, S. N., Cunningham, F., Hung, A. W., Silvestre, H. L., Wen, S., Blundell, T. L., Abell, C., and McConkey, G. A. (2014) Pantothenic acid biosynthesis in the parasite Toxoplasma gondii: a target for chemotherapy. *Antimicrob. Agents Chemother.* **58**, 6345–53

83. Gerdes, S. Y., Scholle, M. D., D'Souza, M., Bernal, A., Baev, M. V, Farrell, M., Kurnasov, O. V, Daugherty, M. D., Mseeh, F., Polanuyer, B. M., Campbell, J. W., Anantha, S., Shatalin, K. Y., Chowdhury, S. A. K., Fonstein, M. Y., and Osterman, A. L. (2002) From genetic footprinting to antimicrobial drug targets: examples in cofactor biosynthetic pathways. *J. Bacteriol.* **184**, 4555–72

84. Vallari, D. S., and Rock, C. O. (1985) Pantothenate transport in Escherichia coli. *J. Bacteriol.* **162**, 1156–61

85. Jackowski, S., and Alix, J. H. (1990) Cloning, sequence, and expression of the pantothenate permease (panF) gene of Escherichia coli. *J. Bacteriol.* **172**, 3842–8

86. Nakamura, H., and Tamura, Z. (1973) Pantothenate uptake in Escherichia coli K-12. *J. Nutr. Sci. Vitaminol. (Tokyo).* **19**, 389–400

87. Reizer, J., Reizer, A., and Saier, M. H. (1990) The Na+/pantothenate symporter (PanF) of Escherichia coli is homologous to the Na+/proline symporter (PutP) of E. coli and the Na+/glucose symporters of mammals. *Res. Microbiol.* **141**, 1069–72

88. Stolz, J., and Sauer, N. (1999) The fenpropimorph resistance gene FEN2 from Saccharomyces cerevisiae encodes a plasma membrane H+-pantothenate symporter. *J. Biol. Chem.* **274**, 18747–52

89. Stolz, J., Caspari, T., Carr, A. M., and Sauer, N. (2004) Cell division defects of Schizosaccharomyces pombe liz1- mutants are caused by defects in pantothenate uptake. *Eukaryot. Cell*. **3**, 406–12

90. Jaehme, M., and Slotboom, D. J. (2015) Diversity of membrane transport proteins for vitamins in bacteria and archaea. *Biochim. Biophys. Acta - Gen. Subj.* **1850**, 565–76

91. Rodionov, D. A., Hebbeln, P., Eudes, A., ter Beek, J., Rodionova, I. A., Erkens, G. B., Slotboom, D. J., Gelfand, M. S., Osterman, A. L., Hanson, A. D., and Eitinger, T. (2009) A novel class of modular transporters for vitamins in prokaryotes. *J. Bacteriol.* **191**, 42–51

92. Neubauer, O., Alfandega, A., Schoknecht, J., Sternberg, U., Pohlmann, A., and Eitinger, T. (2009) Two essential arginine residues in the T components of energy-coupling factor transporters. *J. Bacteriol.* **191**, 6482–8

93. Jackowski, S., and Rock, C. O. (1984) Metabolism of 4'-phosphopantetheine in Escherichia coli. *J. Bacteriol.* **158**, 115–20

94. Jackowski, S., and Rock, C. O. (1984) Turnover of the 4'-phosphopantetheine prosthetic group of acyl carrier protein. *J. Biol. Chem.* **259**, 1891–5

95. Vallari, D. S., and Jackowski, S. (1988) Biosynthesis and degradation both contribute to the regulation of coenzyme A content in Escherichia coli. *J. Bacteriol.* **170**, 3961–6

96. Strauss, E., and Begley, T. P. (2002) The antibiotic activity of N-pentylpantothenamide results from its conversion to ethyldethia-coenzyme a, a coenzyme a antimetabolite. *J. Bol. Chem.* **277**, 48205–9

97. Clarke, K. M., Mercer, A. C., La Clair, J. J., and Burkart, M. D. (2005) In Vivo Reporter Labeling of Proteins via Metabolic Delivery of Coenzyme A Analogues. *J. Am. Chem. Soc.* **127**, 11234–5

98. Meier, J. L., Mercer, A. C., Rivera, H., and Burkart, M. D. (2006) Synthesis and Evaluation of Bioorthogonal Pantetheine Analogues for in Vivo Protein Modification. *J. Am. Chem. Soc.* **128**, 12174–84

99. Craig, J. A., and Snell, E. E. (1951) The comparative activities of pantethine, pantothenic acid, and coenzyme A for various microorganisms. *J. Bacteriol.* **61**, 283–91

100. Dunn, S. D., and Snell, E. E. (1979) Isolation of temperature-sensitive pantothenate kinase mutants of Salmonella typhimurium and mapping of the coaA gene. *J. Bacteriol.* **140**, 805–8

101. Song, W. J., and Jackowski, S. (1992) Cloning, sequencing, and expression of the pantothenate kinase (coaA) gene of Escherichia coli. *J. Bacteriol.* **174**, 6411–7

102. Strauss, E., Kinsland, C., Ge, Y., McLafferty, F. W., and Begley, T. P. (2001) Phosphopantothenoylcysteine synthetase from Escherichia coli. Identification and characterization of the last unidentified coenzyme A biosynthetic enzyme in bacteria. *J. Bol. Chem.* **276**, 13513–6

103. Yang, K., Eyobo, Y., Brand, L. A., Martynowski, D., Tomchick, D., Strauss, E., and Zhang, H. (2006) Crystal structure of a type III pantothenate kinase: insight into the mechanism of an essential coenzyme A biosynthetic enzyme universally distributed in bacteria. *J. Bacteriol.* **188**, 5532–40

104. Awasthy, D., Ambady, A., Bhat, J., Sheikh, G., Ravishankar, S., Subbulakshmi, V., Mukherjee, K., Sambandamurthy, V., and Sharma, U. (2010) Essentiality and functional analysis of type I and type III pantothenate kinases of Mycobacterium tuberculosis. *Microbiology*. **156**, 2691–2701

105. Vallari, D. S., and Rock, C. O. (1987) Isolation and characterization of temperature-sensitive pantothenate kinase (coaA) mutants of Escherichia coli. *J. Bacteriol.* **169**, 5795–800

106. Song, W., and Jackowski, S. (1994) Kinetics and regulation of pantothenate kinase from Escherichia coli. *J. Biol. Chem.* **269**, 27051–8

107. Cheek, S., Ginalski, K., Zhang, H., and Grishin, N. V (2005) A comprehensive update of the sequence and structure classification of kinases. *BMC Struct. Biol.* **5**, 6

108. Cheek, S., Zhang, H., and Grishin, N. V. (2002) Sequence and Structure Classification of Kinases. *J. Mol. Biol.* **320**, 855–81

109. Vallari, D. S., Jackowski, S., and Rock, C. O. (1987) Regulation of pantothenate kinase by coenzyme A and its thioesters. *J. Biol. Chem.* **262**, 2468–71

110. Walker, J. E., Saraste, M., Runswick, M. J., and Gay, N. J. (1982) Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J.* **1**, 945–51

111. Worthington, A. S., and Burkart, M. D. (2006) One-pot chemo-enzymatic synthesis of reporter-modified proteins. *Org. Biomol. Chem.* **4**, 44–6

112. Tilton, G. B., Wedemeyer, W. J., Browse, J., and Ohlrogge, J. (2006) Plant coenzyme A biosynthesis: characterization of two pantothenate kinases from Arabidopsis. *Plant Mol. Biol.* **61**, 629–42

113. Falk, K. L., and Guerra, D. J. (1993) Coenzyme A Biosynthesis in Plants: Partial Purification and Characterization of Pantothenate Kinase from Spinach. *Arch. Biochem. Biophys.* **301**, 424–30

114. Calder, R. B., Williams, R. S., Ramaswamy, G., Rock, C. O., Campbell, E., Unkles, S. E., Kinghorn, J. R., and Jackowski, S. (1999) Cloning and characterization of a eukaryotic pantothenate kinase gene (panK) from Aspergillus nidulans. *J. Biol. Chem.* **274**, 2014–20

115. Zhang, Y.-M., Rock, C. O., and Jackowski, S. (2005) Feedback regulation of murine pantothenate kinase 3 by coenzyme A and coenzyme A thioesters. *J. Biol. Chem.* **280**, 32594–601

116. Ramaswamy, G., Karim, M. A., Murti, K. G., and Jackowski, S. (2004) PPARalpha controls the intracellular coenzyme A concentration via regulation of PANK1alpha gene expression. *J. Lipid Res.* **45**, 17–31

117. Yang, Y., Wu, Z., Kuo, Y. M., and Zhou, B. (2005) Dietary rescue of fumble—a Drosophila model for pantothenate-kinase-associated neurodegeneration. *J. Inherit. Metab. Dis.* **28**, 1055–64

118. Rock, C. O., Calder, R. B., Karim, M. A., and Jackowski, S. (2000) Pantothenate kinase regulation of the intracellular concentration of coenzyme A. *J. Biol. Chem.* **275**, 1377–83

119. Hortnagel, K., Prokisch, H., and Meitinger, T. (2003) An isoform of hPANK2, deficient in pantothenate kinase-associated neurodegeneration, localizes to mitochondria. *Hum. Mol. Genet.* **12**, 321–7

120. Zhou, B., Westaway, S. K., Levinson, B., Johnson, M. A., Gitschier, J., and Hayflick, S. J. (2001) A novel pantothenate kinase gene (PANK2) is defective in Hallervorden-Spatz syndrome. *Nat. Genet.* **28**, 345–9

121. Ni, X., Ma, Y., Cheng, H., Jiang, M., Ying, K., Xie, Y., and Mao, Y. (2002) Cloning and characterization of a novel human pantothenate kinase gene. *Int. J. Biochem. Cell Biol.* **34**, 109–15

122. Rock, C. O., Karim, M. A., Zhang, Y.-M., and Jackowski, S. (2002) The murine pantothenate kinase (Pank1) gene encodes two differentially regulated pantothenate kinase isozymes. *Gene*. **291**, 35–43

123. Leonardi, R., Rock, C. O., Jackowski, S., and Zhang, Y.-M. (2007) Activation of human mitochondrial pantothenate kinase 2 by palmitoylcarnitine. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 1494–9

124. Lehane, A. M., Marchetti, R. V, Spry, C., van Schalkwyk, D. A., Teng, R., Kirk, K., and Saliba, K. J. (2007) Feedback inhibition of pantothenate kinase regulates pantothenol uptake by the malaria parasite. *J. Biol. Chem.* **282**, 25395–405

125. Virga, K. G., Zhang, Y.-M., Leonardi, R., Ivey, R. A., Hevener, K., Park, H.-W., Jackowski, S., Rock, C. O., and Lee, R. E. (2006) Structure–activity relationships and enzyme inhibition of pantothenamide-type pantothenate kinase inhibitors. *Bioorg. Med. Chem.* **14**, 1007–20

126. Hong, B. S., Senisterra, G., Rabeh, W. M., Vedadi, M., Leonardi, R., Zhang, Y.-M., Rock, C. O., Jackowski, S., and Park, H.-W. (2007) Crystal structures of human pantothenate kinases: Insights into allosteric regulation and mutations linked to a neurodegeneration disorder. *J. Biol. Chem.* **282**, 27984–93

127. Hong, B. S., Yun, M. K., Zhang, Y.-M., Chohnan, S., Rock, C. O., White, S. W., Jackowski, S., Park, H.-W., and Leonardi, R. (2006) Prokaryotic Type II and Type III Pantothenate Kinases: The Same Monomer Fold Creates Dimers with Distinct Catalytic Properties. *Structure*. **14**, 1251–61

128. Choudhry, A. E., Mandichak, T. L., Broskey, J. P., Egolf, R. W., Kinsland, C., Begley, T. P., Seefeld, M. A., Ku, T. W., Brown, J. R., Zalacain, M., and Ratnam, K. (2003) Inhibitors of pantothenate kinase: Novel antibiotics for staphylococcal infections. *Antimicrob. Agents Chemother.* **47**, 2051–5

129. Leonardi, R., Chohnan, S., Zhang, Y.-M., Virga, K. G., Lee, R. E., Rock, C. O., and Jackowski, S. (2005) A pantothenate kinase from Staphylococcus aureus refractory to feedback regulation by coenzyme A. *J. Biol. Chem.* **280**, 3314–22

130. Osterman, A., and Overbeek, R. (2003) Missing genes in metabolic pathways: a comparative genomics approach. *Curr. Opin. Chem. Biol.* **7**, 238–51

131. Yocum, R. R., and Patterson, T. A. (2002) Microorganisms and assays for the identification of antibiotics acting on the pantothenate kinase encoded by the coaX gene. USA Patent.

132. Yang, K., Strauss, E., Huerta, C., and Zhang, H. (2008) Structural Basis for Substrate Binding and the Catalytic Mechanism of Type III Pantothenate Kinase. *Biochemistry*. **47**, 1369–80

133. Brand, L. A., and Strauss, E. (2005) Characterization of a new pantothenate kinase isoform from Helicobacter pylori. *J. Biol. Chem.* **280**, 20185–8

134. Kupke, T., Hernandez-Acosta, P., and Culianez-Macia, F. A. (2003) 4'-Phosphopantetheine and coenzyme A biosynthesis in plants. *J. Biol. Chem.* **278**, 38229–37

135. Daugherty, M., Polanuyer, B., Farrell, M., Scholle, M., Lykidis, A., de Crécy-Lagard, V., and Osterman, A. (2002) Complete reconstitution of the human coenzyme A biosynthetic pathway via comparative genomics. *J. Biol. Chem.* **277**, 21431–9

136. Abiko, Y. (1970) [57] Pantothenic acid and coenzyme A: Phosphopantothenoylcysteine synthetase from rat liver (pantothenate 4′-phosphate:l-cysteine ligase, EC 6.3.2.5). *Methods Enzymol.* **18**, 350–4

137. Stanitzek, S., Augustin, M. A., Huber, R., Kupke, T., and Steinbacher, S. (2004) Structural Basis of CTP-Dependent Peptide Bond Formation in Coenzyme A Biosynthesis Catalyzed by Escherichia coli PPC Synthetase. *Structure*. **12**, 1977–88

138. Strauss, E. (2003) *Thiols, radicals and antibiotics: Mechanistic studies in coenzyme A biosynthesis*. PhD thesis, Cornell Uni.

139. Strauss, E., and Begley, T. P. (2001) Mechanistic Studies on Phosphopantothenoylcysteine Decarboxylase. *J. Am. Chem. Soc.* **123**, 6449–50

140. Strauss, E., and Begley, T. P. (2005) The Selectivity for Cysteine over Serine in Coenzyme A Biosynthesis. *ChemBioChem*. **6**, 284–86

141. Stewart, C. J., and Ball, W. J. (1966) Coenzyme A Analogs. II. Enzymatic Conversion of D-Oxypantetheine 4'-Phosphate to Oxy-Coenzyme A. *Biochemistry.* **5**, 3883–6

142. Blaesse, M., Kupke, T., Huber, R., and Steinbacher, S. (2000) Crystal structure of the peptidyl-cysteine decarboxylase EpiD complexed with a pentapeptide substrate. *EMBO J.* **19**, 6299–310

143. Kupke, T., Uebele, M., Schmid, D., Jung, G., Blaesse, M., and Steinbacher, S. (2000) Molecular characterization of lantibiotic-synthesizing enzyme EpiD reveals a function for bacterial Dfp proteins in coenzyme A biosynthesis. *J. Biol. Chem.* **275**, 31838–46

144. Steinbacher, S., Hernandez-Acosta, P., Bieseler, B., Blaesse, M., Huber, R., Culianez-Macia, F. A., and Kupke, T. (2003) Crystal Structure of the Plant PPC Decarboxylase AtHAL3a Complexed with an Ene-thiol Reaction Intermediate. *J. Mol. Biol.* **327**, 193–202

145. Kupke, T., and Schwarz, W. (2006) 4'-phosphopantetheine biosynthesis in Archaea. *J. Biol. Chem.* **281**, 5435–44

146. Martin, D. P., and Drueckhammer, D. G. (1993) Separate Enzymes Catalyze the Final Two Steps of Coenzyme A Biosynthesis in Brevibacterium ammoniagenes: Purification of Pantetheine Phosphate Adenylyltransferase. *Biochem. Biophys. Res. Commun.* **192**, 1155–61

147. Geerlof, A., Lewendon, A., and Shaw, W. V (1999) Purification and characterization of phosphopantetheine adenylyltransferase from Escherichia coli. *J. Biol. Chem.* **274**, 27105–11

148. Izard, T. (2002) The crystal structures of phosphopantetheine adenylyltransferase with bound substrates reveal the enzyme's catalytic mechanism. *J. Mol. Biol.* **315**, 487–95

149. Izard, T. (2003) A novel adenylate binding site confers phosphopantetheine adenylyltransferase interactions with coenzyme A. *J. Bacteriol.* **185**, 4074–80

150. Izard, T., and Geerlof, A. (1999) The crystal structure of a novel bacterial adenylyltransferase reveals half of sites reactivity. *EMBO J.* **18**, 2021–30

151. Miller, J. R., Ohren, J., Sarver, R. W., Mueller, W. T., de Dreu, P., Case, H., and Thanabal, V. (2007) Phosphopantetheine adenylyltransferase from Escherichia coli: investigation of the kinetic mechanism and role in regulation of coenzyme A biosynthesis. *J. Bacteriol.* **189**, 8196–205

152. Worrall, D. M., and Tubbs, P. K. (1983) A bifunctional enzyme complex in coenzyme A biosynthesis: purification of pantetheine phosphate adenylyltransferase and dephospho-CoA kinase. *Biochem. J.* **215**, 153–7

153. Zhyvoloup, A., Nemazanyy, I., Babich, A., Panasyuk, G., Pobigailo, N., Vudmaska, M., Naidenov, V., Kukharenko, O., Palchevskii, S., Savinska, L., Ovcharenko, G., Verdier, F., Valovka, T., Fenton, T., Rebholz, H., Wang, M.-L., Shepherd, P., Matsuka, G., Filonenko, V., and Gout, I. T. (2002) Molecular cloning of CoA Synthase: The missing link in CoA biosynthesis. *J. Biol. Chem.* **277**, 22107–10

154. Aghajanian, S., and Worrall, D. M. (2002) Identification and characterization of the gene encoding the human phosphopantetheine adenylyltransferase and dephospho-CoA kinase bifunctional enzyme (CoA synthase). *Biochem. J.* **365**, 13–8

155. Bork, P., Holm, L., Koonin, E. V., and Sander, C. (1995) The cytidylyltransferase superfamily: Identification of the nucleotide-binding site and fold prediction. *Proteins*. **22**, 259–66

156. Mishra, P., Park, P. K., and Drueckhammer, D. G. (2001) Identification of yacE (coaE) as the structural gene for dephosphocoenzyme A kinase in Escherichia coli K-12. *J. Bacteriol.* **183**, 2774–8

157. O'Toole, N., Barbosa Joao, A. R. G., Li, Y., Hung, L.-W., Matte, A., and Cygler, M. (2003) Crystal structure of a trimeric form of dephosphocoenzyme A kinase from Escherichia coli. *Protein Sci.* **12**, 327–36

158. Schulz, G. E., Muller, C. W., and Diederichs, K. (1990) Induced-fit movements in adenylate kinases. *J. Mol. Biol.* **213**, 627–30

159. Zhang, Y.-M., Frank, M. W., Virga, K. G., Lee, R. E., Rock, C. O., and Jackowski, S. (2004) Acyl carrier protein is a cellular target for the antibacterial action of the pantothenamide class of pantothenate antimetabolites. *J. Biol. Chem.* **279**, 50969–75

160. Björkelid, C., Bergfors, T., Raichurkar, A. K. V, Mukherjee, K., Malolanarasimhan, K., Bandodkar, B., and Jones, T. A. (2013) Structural and biochemical characterization of compounds inhibiting Mycobacterium tuberculosis pantothenate kinase. *J. Biol. Chem.* **288**, 18260–70

161. Reddy, B. K. K., Landge, S., Ravishankar, S., Patil, V., Shinde, V., Tantry, S., Kale, M., Raichurkar, A., Menasinakai, S., Mudugal, N. V., Ambady, A., Ghosh, A., Tunduguru, R., Kaur, P., Singh, R., Kumar, N., Bharath, S., Sundaram, A., Bhat, J., Sambandamurthy, V. K., Björkelid, C., Jones, T. A., Das, K., Bandodkar, B., Malolanarasimhan, K., Mukherjee, K., and Ramachandran, V. (2014) Assessment of Mycobacterium tuberculosis pantothenate kinase vulnerability through target knockdown and mechanistically diverse inhibitors. *Antimicrob. Agents Chemother.* **58**, 3312–26

162. Venkatraman, J., Bhat, J., Solapure, S. M., Sandesh, J., Sarkar, D., Aishwarya, S., Mukherjee, K., Datta, S., Malolanarasimhan, K., Bandodkar, B., and Das, K. S. (2012) Screening, Identification, and Characterization of Mechanistically Diverse Inhibitors of the Mycobacterium Tuberculosis Enzyme, Pantothenate Kinase (CoaA). *J. Biomol. Screen.* **17**, 293–302

163. Rowan, A. S., Nicely, N. I., Cochrane, N., Wlassoff, W. A., Claiborne, A., and Hamilton, C. J. (2009) Nucleoside triphosphate mimicry: a sugar triazolyl nucleoside as an ATP-competitive inhibitor of B. anthracis pantothenate kinase. *Org. Biomol. Chem.* **7**, 4029–36

164. Zhao, L., Allanson, N. M., Thomson, S. P., Maclean, J. K. F., Barker, J. J., Primrose, W. U., Tyler, P. D., and Lewendon, A. (2003) Inhibitors of phosphopantetheine adenylyltransferase. *Eur. J. Med. Chem.* **38**, 345–49

165. Miller, J. R., Thanabal, V., Melnick, M. M., Lall, M., Donovan, C., Sarver, R. W., Lee, D.-Y., Ohren, J., and Emerson, D. (2010) The Use of Biochemical and Biophysical Tools for Triage of High-Throughput Screening Hits - A Case Study with Escherichia coli Phosphopantetheine Adenylyltransferase. *Chem. Biol. Drug Des.* **75**, 444–54

166. de Jonge, B. L. M., Walkup, G. K., Lahiri, S. D., Huynh, H., Neckermann, G., Utley, L., Nash, T. J., Brock, J., San Martin, M., Kutschke, A., Johnstone, M., Laganas, V., Hajec, L., Gu, R.-F., Ni, H., Chen, B., Hutchings, K., Holt, E., McKinney, D., Gao, N., Livchak, S., and Thresher, J. (2013) Discovery of inhibitors of 4'-phosphopantetheine adenylyltransferase (PPAT) to validate PPAT as a target for antibacterial therapy. *Antimicrob. Agents Chemother.* **57**, 6005–15

167. Cheng, C.-S., Jia, K.-F., Chen, T., Chang, S.-Y., Lin, M.-S., and Yin, H.-S. (2013) Experimentally Validated Novel Inhibitors of Helicobacter pylori Phosphopantetheine Adenylyltransferase Discovered by Virtual High-Throughput Screening. *PLoS One*. **8**, e74271

168. Tommasi, R., Brown, D. G., Walkup, G. K., Manchester, J. I., and Miller, A. A. (2015) ESKAPEing the labyrinth of antibacterial discovery. *Nat. Rev. Drug Discov.* **14**, 529–42

169. Arraes, F. B. M., Carvalho, M. J. A. de, Maranhão, A. Q., Brígido, M. M., Pedrosa, F. O., and Felipe, M. S. S. (2007) Differential metabolism of Mycoplasma species as revealed by their genomes. *Genet. Mol. Biol.* **30**, 182–9

170. Hartfield, M., Lía Murall, C., and Alizon, S. (2014) Clinical applications of pathogen phylogenies. *Trends Mol. Med.* **20**, 394–404

171. Brocchieri, L. (2001) Phylogenetic Inferences from Molecular Sequences: Review and Critique. *Theor. Popul. Biol.* **59**, 27–40

172. Schuwirth, B. S., Borovinskaya, M. A., Hau, C. W., Zhang, W., Vila-Sanjurjo, A., Holton, J. M., and Cate, J. H. D. (2005) Structures of the bacterial ribosome at 3.5 A resolution. *Science.* **310**, 827–34

173. Kitahara, K., Yasutake, Y., and Miyazaki, K. (2012) Mutational robustness of 16S ribosomal RNA, shown by experimental horizontal gene transfer in Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 19220–5

174. Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J. D., and Higgins, D. G. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539

175. Sievers, F., and Higgins, D. G. (2017) Clustal Omega for making accurate alignments of many protein sequences. *Protein Sci.* **27**, 135–45

176. Soding, J. (2005) Protein homology detection by HMM-HMM comparison. *Bioinformatics*. **21**, 951–60

177. Blackshields, G., Sievers, F., Shi, W., Wilm, A., and Higgins, D. G. (2010) Sequence embedding for fast construction of guide trees for multiple sequence alignment. *Algorithms Mol. Biol.* **5**, 21

178. Yang, Z., and Rannala, B. (2012) Molecular phylogenetics: principles and practice. *Nat. Rev. Genet.* **13**, 303–14

179. Saitou, N., and Nei, M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**, 406–25

180. Wiley, E., and Lieberman, B. (2011) Parametric phylogenetics. in *Phylogenetics: Theory and Practice of Phylogenetic Systematics*, 2nd Ed., pp. 203–28, John Wiley & Sons, New Jersey, USA

181. Wiley, E., and Lieberman, B. (2011) Parsimony and parsimony analysis. in *Phylogenetics: Theory and Practice of Phylogenetic Systematics*, 2nd Ed., pp. 152–202, John Wiley & Sons, New Jersey, USA

182. Felsenstein, J. (1981) Evolutionary trees from DNA sequences: A maximum likelihood approach. *J. Mol. Evol.* **17**, 368–76

183. Whelan, S., and Morrison, D. A. (2017) Inferring Trees. in *Bioinformatics: Volume I: Data, Sequence Analysis, and Evolution*, 2nd Ed. (Keith, J. M. ed), pp. 349–77, Humana Press, New York, USA

184. Chor, B., and Tuller, T. (2005) Maximum likelihood of evolutionary trees: hardness and approximation. *Bioinformatics*. **21**, i97-106

185. Miller, M., Pfeiffer, W., and Schwartz, T. (2010) Creating the CIPRES Science Gateway for inference of large phylogenetic trees. in *Proceedings of the Gateway Computing Environments Workshop (GCE)*, pp. 1–8, New Orleans, USA

186. Stamatakis, A. (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*. **30**, 1312–13

187. Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., and Morishima, K. (2017) KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* **45**, D353-61

188. Kanehisa, M., Goto, S., Hattori, M., Aoki-Kinoshita, K. F., Itoh, M., Kawashima, S., Katayama, T., Araki, M., and Hirakawa, M. (2006) From genomics to chemical genomics: new developments in KEGG. *Nucleic Acids Res.* **34**, D354-7

189. Kanehisa, M., Goto, S., Furumichi, M., Tanabe, M., and Hirakawa, M. (2010) KEGG for representation and analysis of molecular networks involving diseases and drugs. *Nucleic Acids Res.* **38**, D355-60

190. Overbeek, R., Begley, T., Butler, R. M., Choudhuri, J. V, Chuang, H.-Y., Cohoon, M., de Crécy-Lagard, V., Diaz, N., Disz, T., Edwards, R., Fonstein, M., Frank, E. D., Gerdes, S., Glass, E. M., Goesmann, A., Hanson, A., Iwata-Reuyl, D., Jensen, R., Jamshidi, N., Krause, L., Kubal, M., Larsen, N., Linke, B., McHardy, A. C., Meyer, F., Neuweger, H., Olsen, G., Olson, R., Osterman, A., Portnoy, V., Pusch, G. D., Rodionov, D. A., Rückert, C., Steiner, J., Stevens, R., Thiele, I., Vassieva, O., Ye, Y., Zagnitko, O., and Vonstein, V. (2005) The Subsystems Approach to Genome Annotation and its Use in the Project to Annotate 1000 Genomes. *Nucleic Acids Res.* **33**, 5691–702

191. Pruitt, K. D., Tatusova, T., Brown, G. R., and Maglott, D. R. (2012) NCBI Reference Sequences (RefSeq): current status, new features and genome annotation policy. *Nucleic Acids Res.* **40**, D130-5

192. Pruitt, K. D., Tatusova, T., and Maglott, D. R. (2007) NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res.* **35**, D61-5

193. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–10

194. Johnson, M., Zaretskaya, I., Raytselis, Y., Merezhuk, Y., McGinnis, S., and Madden, T. L. (2008) NCBI BLAST: a better web interface. *Nucleic Acids Res.* **36**, W5-9

195. Altschul, S., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–402

196. Altschul, S. F., and Koonin, E. V (1998) Iterated profile searches with PSI-BLAST--a tool for discovery in protein databases. *Trends Biochem. Sci.* **23**, 444–7

197. Junker, V. L., Apweiler, R., and Bairoch, A. (1999) Representation of functional information in the SWISS-PROT data bank. *Bioinformatics.* **15**, 1066–7

198. Apweiler, R. (2001) Functional information in SWISS-PROT: The basis for large-scale characterisation of protein sequences. *Brief. Bioinform.* **2**, 9–18

199. Kretschmann, E., Fleischmann, W., and Apweiler, R. (2001) Automatic rule generation for protein annotation with the C4.5 data mining algorithm applied on SWISS-PROT. *Bioinformatics*. **17**, 920–6

200. Marchler-Bauer, A., Bo, Y., Han, L., He, J., Lanczycki, C. J., Lu, S., Chitsaz, F., Derbyshire, M. K., Geer, R. C., Gonzales, N. R., Gwadz, M., Hurwitz, D. I., Lu, F., Marchler, G. H., Song, J. S., Thanki, N., Wang, Z., Yamashita, R. A., Zhang, D., Zheng, C., Geer, L. Y., and Bryant, S. H. (2017) CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res.* **45**, D200-3

201. Marchler-Bauer, A., Derbyshire, M. K., Gonzales, N. R., Lu, S., Chitsaz, F., Geer, L. Y., Geer, R. C., He, J., Gwadz, M., Hurwitz, D. I., Lanczycki, C. J., Lu, F., Marchler, G. H., Song, J. S., Thanki, N., Wang, Z., Yamashita, R. A., Zhang, D., Zheng, C., and Bryant, S. H. (2015) CDD: NCBI's conserved domain database. *Nucleic Acids Res.* **43**, D222-6

202. Finn, R. D., Attwood, T. K., Babbitt, P. C., Bateman, A., Bork, P., Bridge, A. J., Chang, H.-Y., Dosztányi, Z., El-Gebali, S., Fraser, M., Gough, J., Haft, D., Holliday, G. L., Huang, H., Huang, X., Letunic, I., Lopez, R., Lu, S., Marchler-Bauer, A., Mi, H., Mistry, J., Natale, D. A., Necci, M., Nuka, G., Orengo, C. A., Park, Y., Pesseat, S., Piovesan, D., Potter, S. C., Rawlings, N. D., Redaschi, N., Richardson, L., Rivoire, C., Sangrador-Vegas, A., Sigrist, C., Sillitoe, I., Smithers, B., Squizzato, S., Sutton, G., Thanki, N., Thomas, P. D., Tosatto, S. C. E., Wu, C. H., Xenarios, I., Yeh, L.-S., Young, S.-Y., and Mitchell, A. L. (2017) InterPro in 2017—beyond protein family and domain annotations. *Nucleic Acids Res.* **45**, D190-9

203. Bailey, T. L., Boden, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., Ren, J., Li, W. W., and Noble, W. S. (2009) MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.* **37**, W202-8

204. Bailey, T. L., Johnson, J., Grant, C. E., and Noble, W. S. (2015) The MEME Suite. *Nucleic Acids Res.* **43**, W39-49

205. Hall, T. A. (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* **41**, 95–98

206. Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., and Drummond, A. (2012) Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*. **28**, 1647–9

207. Nicely, N. I., Parsonage, D., Paige, C., Newton, G. L., Fahey, R. C., Leonardi, R., Jackowski, S., Mallett, T. C., and Claiborne, A. (2007) Structure of the Type III Pantothenate Kinase from Bacillus anthracis at 2.0 Å Resolution: Implications for Coenzyme A-Dependent Redox Biology. *Biochemistry.* **46**, 3234–45

208. Bork, P., Sander, C., and Valencia, A. (1992) An ATPase domain common to prokaryotic cell cycle proteins, sugar kinases, actin, and hsp70 heat shock proteins. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 7290–4

209. Kupke, T. (2001) Molecular characterization of the 4'-phosphopantothenoylcysteine decarboxylase domain of bacterial Dfp flavoproteins. *J. Biol. Chem.* **276**, 27597–604

210. Kupke, T. (2002) Molecular characterization of the 4'-phosphopantothenoylcysteine synthetase domain of bacterial dfp flavoproteins. *J. Biol. Chem.* **277**, 36137–45

211. Lawhorn, B. G., Gerdes, S. Y., and Begley, T. P. (2004) A genetic screen for the identification of thiamin metabolic genes. *J. Biol. Chem.* **279**, 43555–9

212. Kuznetsova, E., Proudfoot, M., Gonzalez, C. F., Brown, G., Omelchenko, M. V, Borozan, I., Carmel, L., Wolf, Y. I., Mori, H., Savchenko, A. V, Arrowsmith, C. H., Koonin, E. V, Edwards, A. M., and Yakunin, A. F. (2006) Genome-wide analysis of substrate specificities of the Escherichia coli haloacid dehalogenase-like phosphatase family. *J. Biol. Chem.* **281**, 36149–61

213. Messick, J. B. (2004) Hemotrophic mycoplasmas (hemoplasmas): a review and new insights into pathogenic potential. *Vet. Clin. Pathol.* **33**, 2–13

214. Guimaraes, A. M. S., Santos, A. P., do Nascimento, N. C., Timenetsky, J., and Messick, J. B. (2014) Comparative Genomics and Phylogenomics of Hemotrophic Mycoplasmas. *PLoS One*. **9**, e91445

215. Hutchison, C. A., Chuang, R.-Y., Noskov, V. N., Assad-Garcia, N., Deerinck, T. J., Ellisman, M. H., Gill, J., Kannan, K., Karas, B. J., Ma, L., Pelletier, J. F., Qi, Z.-Q., Richter, R. A., Strychalski, E. A., Sun, L., Suzuki, Y., Tsvetanova, B., Wise, K. S., Smith, H. O., Glass, J. I., Merryman, C., Gibson, D. G., and Venter, J. C. (2016) Design and synthesis of a minimal bacterial genome. *Science.* **351**, aad6253

216. Abolnik, C. (2015) Complete Genome Sequence of Mycoplasma meleagridis, a Possible Emerging Pathogen in Chickens. *Genome Announc.* **3**, e00380-15

217. Hadley, C. (2003) Righting the wrongs. *EMBO Rep.* **4**, 829–31

218. Denton, J. F., Lugo-Martinez, J., Tucker, A. E., Schrider, D. R., Warren, W. C., and Hahn, M. W. (2014) Extensive error in the number of genes inferred from draft genome assemblies. *PLoS Comput. Biol.* **10**, e1003998

219. Driscoll, T. P., Verhoeve, V. I., Guillotte, M. L., Lehman, S. S., Rennoll, S. A., Beier-Sexton, M., Rahman, M. S., Azad, A. F., and Gillespie, J. J. (2017) Wholly Rickettsia! Reconstructed Metabolic Profile of the Quintessential Bacterial Parasite of Eukaryotic Cells. *MBio*. **8**, e00859-17

220. Großhennig, S., Schmidl, S. R., Schmeisky, G., Busse, J., and Stülke, J. (2013) Implication of glycerol and phospholipid transporters in Mycoplasma pneumoniae growth and virulence. *Infect. Immun.* **81**, 896–904

221. Vozza, A., De Leonardis, F., Paradies, E., De Grassi, A., Pierri, C. L., Parisi, G., Marobbio, C. M. T., Lasorsa, F. M., Muto, L., Capobianco, L., and Dolce, V. (2017) Biochemical characterization of a new mitochondrial transporter of dephosphocoenzyme A in Drosophila melanogaster. *Biochim. Biophys. Acta - Bioenerg.* **1858**, 137–46

222. Sibon, O., and Strauss, E. (2016) Coenzyme A: to make it or uptake it? *Nat. Rev. Mol. Cell Biol*. **17**, 605–6

223. Srinivasan, B., Baratashvili, M., van der Zwaag, M., Kanon, B., Colombelli, C., Lambrechts, R. A., Schaap, O., Nollen, E. A., Podgoršek, A., Kosec, G., Petković, H., Hayflick, S., Tiranti, V., Reijngoud, D.-J., Grzeschik, N. A., and Sibon, O. C. M. (2015) Extracellular 4′-phosphopantetheine is a source for intracellular coenzyme A synthesis. *Nat. Chem. Biol.* **11**, 784–92

224. Wei, S., Guo, Z., Li, T., Zhang, T., Li, X., Zhou, Z., Li, Z., Liu, M., Luo, R., Bi, D., Chen, H., Zhou, R., and Jin, H. (2012) Genome sequence of Mycoplasma iowae strain 695, an unusual pathogen causing deaths in turkeys. *J. Bacteriol.* **194**, 547–8

225. Berry, V., and Gascuel, O. (1996) On the interpretation of bootstrap trees: appropriate threshold of clade selection and induced gain. *Mol. Biol. Evol.* **13**, 999–1011

226. Halbedel, S., and Stülke, J. (2007) Tools for the genetic analysis of Mycoplasma. *Int. J. Med. Microbiol.* **297**, 37–44

227. Smiley, B. K., and Minion, F. C. (1993) Enhanced readthrough of opal (UGA) stop codons and production of Mycoplasma pneumoniae P1 epitopes in Escherichia coli. *Gene*. **134**, 33–40

228. Stamburski, C., Renaudin, J., and Bove, J. M. (1991) First step toward a virus-derived vector for gene cloning and expression in spiroplasmas, organisms which read UGA as a tryptophan codon: synthesis of chloramphenicol acetyltransferase in Spiroplasma citri. *J. Bacteriol.* **173**, 2225–30

229. Ishii, T. M., Zerr, P., Xia, X. M., Bond, C. T., Maylie, J., and Adelman, J. P. (1998) Site-directed mutagenesis. *Methods Enzymol.* **293**, 53–71

230. Dieffenbach, C. W., Lowe, T. M., and Dveksler, G. S. (1993) General concepts for PCR primer design. *Genome Res.* **3**, S30-7

231. Goosen, R., and Strauss, E. (2017) Simultaneous quantification of coenzyme A and its salvage pathway intermediates in in vitro and whole cell-sourced samples. *RSC Adv.* **7**, 19717–24

232. Eschenfeldt, W. H., Maltseva, N., Stols, L., Donnelly, M. I., Gu, M., Nocek, B., Tan, K., Kim, Y., and Joachimiak, A. (2010) Cleavable C-terminal His-tag vectors for structure determination. *J. Struct. Funct. Genomics*. **11**, 31–9

233. Zhao, D., and Huang, Z. (2016) Effect of His-Tag on Expression, Purification, and Structure of Zinc Finger Protein, ZNF191(243-368). *Bioinorg. Chem. Appl.* **2016**, 1–6

234. Kapust, R. B., and Waugh, D. S. (1999) Escherichia coli maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Sci.* **8**, 1668–74

235. Nygren, P.-Å., Stahl, S., and Uhlén, M. (1994) Engineering proteins to facilitate bioprocessing. *Trends Biotechnol.* **12**, 184–8

236. Hammarström, M., Hellgren, N., van den Berg, S., Berglund, H., and Härd, T. (2002) Rapid screening for improved solubility of small human proteins produced as fusion proteins in Escherichia coli. *Protein Sci.* **11**, 313–21

237. Dyson, M. R., Shadbolt, S. P., Vincent, K. J., Perera, R. L., and McCafferty, J. (2004) Production of soluble mammalian proteins in Escherichia coli: identification of protein features that correlate with successful expression. *BMC Biotechnol.* **4**, 32

238. Busso, D., Delagoutte-Busso, B., and Moras, D. (2005) Construction of a set Gateway-based destination vectors for high-throughput cloning and expression screening in Escherichia coli. *Anal. Biochem.* **343**, 313–21

239. Kataeva, I., Chang, J., Xu, H., Luan, C. H., Zhou, J., Uversky, V. N., Lin, D., Horanyi, P., Liu, Z. J., Ljungdahl, L. G., and Rose, J. (2005) Improving Solubility of Shewanella oneidensis MR-1 and Clostridium thermocellum JW-20 Proteins Expressed into Esherichia coli. *J. Proteome Res.* **4**, 1942–51

240. Braud, S., Moutiez, M., Belin, P., Abello, N., Drevet, P., Zinn-Justin, S., Courçon, M., Masson, C., Dassa, J., Charbonnier, J. B., and Boulain, J. C. (2005) Dual expression system suitable for high-throughput fluorescence-based screening and production of soluble proteins. *J. Proteome Res.* **4**, 2137–47

241. Waugh, D. S. (2004) *MBP Fusion FAQ*, National Cancer Institute

242. Muench, S. P., Rafferty, J. B., Mcleod, R., Rice, D. W., and Prigge, S. T. (2003) Expression, purification and crystallization of the Plasmodium falciparum enoyl reductase. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **59**, 1246–8

243. Kapust, R. B., and Waugh, D. S. (2000) Controlled Intracellular Processing of Fusion Proteins by TEV Protease. *Protein Expr. Purif.* **19**, 312–8

244. Du, Y., Gisselberg, J. E., Johnson, J. D., Lee, P. J., Prigge, S. T., and Bachmann, B. O. (2010) Lactococcus lactis fabH, encoding beta-ketoacyl-acyl carrier protein synthase, can be functionally replaced by the Plasmodium falciparum congener. *Appl. Environ. Microbiol.* **76**, 3959–66

245. Thomas, J. G., Ayling, A., and Baneyx, F. (1997) Molecular chaperones, folding catalysts, and the recovery of active recombinant proteins from E. coli. *Appl. Biochem. Biotechnol.* **66**, 197–238

246. Carrió, M. M., and Villaverde, A. (2001) Protein aggregation as bacterial inclusion bodies is reversible. *FEBS Lett.* **489**, 29–33

247. de Marco, A., and De Marco, V. (2004) Bacteria co-transformed with recombinant proteins and chaperones cloned in independent plasmids are suitable for expression tuning. *J. Biotechnol.* **109**, 45–52

248. Nishihara, K., Kanemori, M., Kitagawa, M., Yanagi, H., and Yura, T. (1998) Chaperone coexpression plasmids: differential and synergistic roles of DnaK-DnaJ-GrpE and GroEL-GroES in assisting folding of an allergen of Japanese cedar pollen, Cryj2, in Escherichia coli. *Appl. Environ. Microbiol.* **64**, 1694–9

249. Wipperman, M. F., Yang, M., Thomas, S. T., and Sampson, N. S. (2013) Shrinking the FadE proteome of Mycobacterium tuberculosis: insights into cholesterol metabolism through identification of an α2β2 heterotetrameric acyl coenzyme A dehydrogenase family. *J. Bacteriol.* **195**, 4331–41

250. Filip, C., Fletcher, G., Wulff, J. L., and Earhart, C. F. (1973) Solubilization of the cytoplasmic membrane of Escherichia coli by the ionic detergent sodium-lauryl sarcosinate. *J. Bacteriol.* **115**, 717–22

251. Frankel, S., Sohn, R., and Leinwand, L. (1991) The use of sarkosyl in generating soluble protein after bacterial expression. *Proc. Natl. Acad. Sci. U. S. A.* **88**, 1192–6

252. Massiah, M. A., Wright, K. M., and Du, H. (2016) Obtaining Soluble Folded Proteins from Inclusion Bodies Using Sarkosyl, Triton X-100, and CHAPS: Application to LB and M9 Minimal Media. *Curr. Protoc. Protein Sci.* **84**, 6.13.1-24

253. Studier, F. W. (2005) Protein production by auto-induction in high-density shaking cultures. *Protein Expr. Purif.* **41**, 207–34

254. Baneyx, F., and Mujacic, M. (2004) Recombinant protein folding and misfolding in Escherichia coli. *Nat. Biotechnol.* **22**, 1399–1408

255. Sørensen, H., and Mortensen, K. (2005) Soluble expression of recombinant proteins in the cytoplasm of Escherichia coli. *Microb. Cell Fact.* **4**, 1–8

256. Ledent, P., Duez, C., Vanhove, M., Lejeune, A., Fonzé, E., Charlier, P., Rhazi-Filali, F., Thamm, I., Guillaume, G., Samyn, B., Devreese, B., Van Beeumen, J., Lamotte-Brasseur, J., and Frère, J.-M. (1997) Unexpected influence of a C-terminal-fused His-tag on the processing of an enzyme and on the kinetic and folding parameters. *FEBS Lett.* **413**, 194–6

257. Wu, J., and Filutowicz, M. (1999) Hexahistidine (His6)-tag dependent protein dimerization: a cautionary tale. *Acta Biochim. Pol.* **46**, 591–9

258. Deutscher, M. P. (2009) Maintaining Protein Stability. in *Methods In Enzymology*, 2nd Ed. (Burgess, R. R., and Deutscher, M. P. eds), pp. 121–5, Academic Press, San Diego, USA

The Supplementary Data of the Thesis:

Bioinformatic characterisation of genes associated with coenzyme A biosynthesis in mycoplasmas and expression and isolation of dephospho-coenzyme A kinase in *Mycoplasma* sp. Ms02

by
Tertius Alwyn Ras

Thesis presented in fulfilment of the requirements for the degree of Master of Science (Biochemistry) in the Faculty of Science at Stellenbosch University

Supervisor: Dr A. Botes
Co-Supervisor: Prof. E. Strauss

March 2018

Table of Contents

Appendix 1 – Supplementary tables and figures for Chapter 3

Supplementary Table 1.1 The 62 *Mycoplasma* species and the accession numbers for their *16S rRNA* nucleotide and respective amino acid sequences

***16S rRNA* gene**		**PanK Type III protein**		**CoaBC protein**		**PPAT protein**		**DPCK protein**	
***Mycoplasma* species**	**NCBI accession number**	***Mycoplasma* species**	**NCBI accession number**	***Mycoplasma* species**	**NCBI accession number**	***Mycoplasma* species**	**NCBI accession number**	***Mycoplasma* species**	**NCBI accession number**
M. agalactiae	NR_118811.1					*M. agalactiae* PG2	WP_011949321.1	*M. agalactiae* PG2	WP_011949819.1
M. alligatoris	NR_041844.1	*M. alligatoris* A21JP2	WP_005683755.1			*M. alligatoris* A21JP2	WP_005683206.1	*M. alligatoris* A21JP2	WP_005683611.1[a]
M. alvi	NR_025985.1	*M. alvi* ATCC 29626	WP_033159884.1			*M. alvi* ATCC 29626	WP_033159881.1	*M. alvi* ATCC 29626	WP_033159717.1
M. anatis	NR_113689.1	*M. anatis* 1340	WP_006886261.1	*M. anatis* 1340	WP_006886260.1	*M. anatis* 1340	WP_040544544.1	*M. anatis* 1340	WP_006886900.1
M. arginini	NR_041743.1	*M. arginini* ATCC 23838	WP_020003050.1	*M. arginini* ATCC 23838	WP_020003049.1	*M. arginini* ATCC 23838	WP_011283787.1	*M. arginini* ATCC 23838	WP_020002982.1
M. arthritidis	NR_113688.1								
M. bovigenitalium	NR_113690.1					*M. bovigenitalium* 51080	WP_004419352.1	*M. bovigenitalium* 51080	WP_004420621.1
M. bovis	NR_102850.1					*M. bovis* 08M	WP_013456618.1	*M. bovis* 08M	WP_014829984.1
M. bovoculi	NR_121731.1							*M. bovoculi* M165/69	WP_022935460.1 (HAD)[c]
M. buteonis	NR_025177.1	*M. buteonis* ATCC 51371	WP_036452417.1			*M. buteonis* ATCC 51371	WP_036452642.1	*M. buteonis* ATCC 51371	WP_036451916.1[a]
M. californicum	NR_029166.1					*M. californicum* ST-6	WP_038561568.1	*M. californicum* ST-6	WP_038561049.1
M. canadense	NR_025988.1							*M. bovigenitalium* 51080	WP_004420621.1
M. canis	NR_113676.1					*M. canis* UF33	WP_004796179.1	*M. canis* UF33	WP_004796406.1[a]
M. capricolum subsp. *capricolum*	NR_074664.1					*M. capricolum*	WP_011387120.1	*M. capricolum subsp. capricolum* ATCC 27343	ABC01831.1
M. capricolum subsp. *capripneumoniae*	NR_118795.1							*M. capricolum subsp. capripneumoniae*	WP_045076023.1
M. collis	NR_114636.1					*M. collis* ATCC 35278	WP_033161390.1	*M. collis* ATCC 35278	WP_033160863.1
M. columbinum	NR_025063.1					*M. columbinum* ATCC 29257	WP_006608789.1	*M. columbinum* ATCC 29257	WP_029891912.1
M. columborale	NR_025179.1	*M. columborale* ATCC 29258	WP_036434680.1	*M. columborale* ATCC 29258	WP_036434682.1	*M. columborale* ATCC 29258	WP_036434731.1	*M. columborale* ATCC 29258	WP_036434144.1

***16S rRNA* gene**		**PanK Type III protein**		**CoaBC protein**		**PPAT protein**		**DPCK protein**	
***Mycoplasma* species**	**NCBI accession number**	***Mycoplasma* species**	**NCBI accession number**	***Mycoplasma* species**	**NCBI accession number**	***Mycoplasma* species**	**NCBI accession number**	***Mycoplasma* species**	**NCBI accession number**
M. conjunctivae	NR_044781.1					*M. conjunctivae* HRC/581T	WP_041594472.1	*M. conjunctivae* HRC/581T	WP_012751785.1 (HAD)[c]
M. cricetuli	NR_025180.1	*M. cricetuli* ATCC 35279	WP_025755140.1	*M. cricetuli* ATCC 35279	WP_025755141.1	*M. cricetuli* ATCC 35279	WP_025755127.1	*M. cricetuli* ATCC 35279	WP_025755469.1[a]
M. crocodyli	NR_074301.1	*M. crocodyli* MP145	WP_013054512.1			*M. crocodyli* MP145	WP_013054470.1	*M. crocodyli* MP145	WP_013054324.1
M. dispar	NR_025182.1							*M. dispar* ATCC 27140	WP_044635407.1 (HAD)[c]
M. felifaucium	NR_025963.1					*M. felifaucium* ATCC 43428	WP_027334914.1	*M. felifaucium* ATCC 43428	WP_027334523.1
M. felis	NR_029174.1					*M. felis* ATCC 23391	WP_036430302.1	*M. felis* ATCC 23391	WP_036430513.1
M. fermentans	NR_113683.1					*M. fermentans* PG18	WP_013526711.1	*M. fermentans* PG18	WP_013354344.1
M. flocculare	NR_036954.1							*M. flocculare* ATCC 27399	WP_002557681.1 (HAD)[ac]
M. gallinaceum	NR_025913.1	*M. gallinaceum* B2096 8B	AKA50150.1[a]	*M. gallinaceum* B2096 8B	AKA50149.1[a]	*M. gallinaceum* B2096 8B	AKA50018.1		
M. gallinarum	NR_113687.1					*M. gallinarum* Mgn_IPT	WP_063626090.1	*M. gallinarum* Mgn_IPT	WP_063626074.1
M. gallisepticum	NR_104952.1							*M. gallisepticum* str. R(low)	WP_011113477.1
M. genitalium	NR_074611.1							*M. genitalium* G37	WP_009885895.1
M. haemocanis	AY529641.1								
M. hominis	NR_041881.1								
M. hyopneumoniae	NR_121689.1							*M. hyopneumoniae* J	WP_011284119.1 (HAD)[c]
M. hyorhinis	NR_041845.1							*M. hyorhinis* HUB-1	WP_041363606.1
M. imitans	NR_025912.1							*M. imitans* ATCC 51306	WP_027121942.1
M. iners	NR_025064.1					*M. iners* ATCC 19705	WP_029512930.1	*M. iners* ATCC 19705	WP_029512864.1
M. iowae	NR_044669.2	*M. iowae* 695	WP_004025380.1	*M. iowae* 695	WP_004025332.1 (PPCS) WP_004025331.1 (PPCDC)	*M. iowae* 695	WP_004024713.1	*M. iowae* 695	WP_036452307.1
M. leachii	NR_044773.1					*M. leachii* PG50	WP_013447590.1	*M. leachii* PG50	WP_013447415.1
M. leonicaptivi	NR_025965.1					*M. leonicaptivi* ATCC 49890	WP_027121244.1	*M. leonicaptivi* ATCC 49890	WP_051521842.1
M. lipofaciens	NR_025065.1					*M. lipofaciens* ATCC 35015	WP_027120410.1	*M. lipofaciens* ATCC 35015	WP_027120753.1

***16S rRNA* gene**		**PanK Type III protein**		**CoaBC protein**		**PPAT protein**		**DPCK protein**	
***Mycoplasma* species**	**NCBI accession number**	***Mycoplasma* species**	**NCBI accession number**	***Mycoplasma* species**	**NCBI accession number**	***Mycoplasma* species**	**NCBI accession number**	***Mycoplasma* species**	**NCBI accession number**
M. mobile	NR_074620.1	*M. mobile* 163K	WP_011265111.1	*M. mobile* 163K	WP_011265112.1	*M. mobile* 163K	WP_011265096.1[b]	*M. mobile* 163K	WP_011265097.1
M. molare	NR_114637.1	*M. molare* ATCC 27746	WP_027123244.1			*M. molare* ATCC 27746	WP_027123383.1	*M. molare* ATCC 27746	WP_027123123.1
M. mycoides subsp. *capri* LC	NR_118794.1					*M. mycoides* subsp. *capri* LC str. 95010	WP_013729412.1	*M. mycoides* subsp. *capri* LC str. 95010	WP_013729251.1
M. mycoides subsp. *mycoides* SC	NR_074703.1					*M. mycoides* subsp. *mycoides* SC str. PG1	NP_975271.1	*M. mycoides* subsp. *mycoides* SC str. PG1	NP_975106.1
M. opalescens	NR_025067.1					*M. opalescens* ATCC 27921	WP_029906491.1	*M. opalescens* ATCC 27921	WP_029906184.1
M. ovipneumoniae	NR_025989.1							*M. ovipneumoniae* SC01	WP_010321428.1 (HAD)[c]
M. ovis	AF338268.1								
M. parvum	AB610850.1								
M. penetrans	NR_118664.1	*M. penetrans* HF-2	WP_011077834.1			*M. penetrans* HF-2	WP_011077780.1	*M. penetrans* HF-2	WP_011077028.1
M. pirum	NR_029165.1	*M. pirum* MPI25960	WP_027124235.1			*M. pirum* MPI25960	WP_027124232.1	*M. pirum* MPI25960	WP_052663051.1
M. pneumoniae	NR_041751.1							*M. pneumoniae* M129	NP_110070.1
M. primatum	NR_025068.1					*M. primatum* ATCC 25948	WP_029513805.1	*M. primatum* ATCC 25948	WP_029513151.1
M. pulmonis	NR_041744.1	*M. pulmonis* UAB CTIP	WP_010925274.1			*M. pulmonis* UAB CTIP	WP_010924901.1	*M. pulmonis* UAB CTIP	WP_010925522.1
M. putrefaciens	NR_025971.1					*M. putrefaciens* KS1	WP_014035258.1	*M. putrefaciens* KS1	WP_014035169.1
M. simbae	NR_025964.1					*M. simbae* ATCC 49888	WP_029608728.1	*M. simbae* ATCC 49888	WP_029608675.1
M. sturni	NR_025968.1	*M. sturni* DSM 22021	WP_036464058.1	*M. sturni* DSM 22021	WP_036464056.1	*M. sturni* DSM 22021	WP_036464160.1	*M. sturni* DSM 22021	WP_036464601.1[a]
M. suis	EU603330.1								
M. synoviae	NR_044811.1	*M. synoviae* 53	WP_041351793.1	*M. synoviae* 53	WP_011283173.1	*M. synoviae* 53	AAZ44058.2[b]	*M. synoviae* 53	WP_041352058.1[a]
M. sp. Ms02	DQ223546.1								
M. testudinis	NR_029175.1	*M. testudinis* ATCC 43263	WP_027120000.1	*M. testudinis* ATCC 43263	WP_084266141.1[a]	*M. testudinis* ATCC 43263	WP_036499102.1	*M. testudinis* ATCC 43263	WP_027119674.1
M. wenyonii	KX171205.1								
M. yeatsii	NR_026037.1					*M. yeatsii* GM274B	WP_042733350.1	*M. yeatsii* GM274B	WP_004427570.1

[a]Annotated as hypothetical protein; [b]Annotated as putative protein; [c]Bifunctional HAD-like/dephospho-coenzyme A kinase protein

Supplementary Table 1.2 The genomic locations of currently annotated CoA biosynthetic pathway enzyme-encoding genes and newly identified homologues within the corresponding *Mycoplasma* genome

***Mycoplasma* species**	**NCBI accession number**	**Genomic location**			
		PanK type III	**CoaBC**	**PPAT**	**DPCK**
M. agalactiae PG2	NC_009497.1^			161810-162232 (locus tag: MAG_RS00720)	complement (784004-784576) (locus tag: MAG_RS03375)
M. alligatoris A21JP2	NZ_ADNC01000000.1*	37520-38281 (locus tag: MALL_RS02820) (contig: NZ_ADNC01000027.1)		complement (40704-41141) (locus tag: MALL_RS00640) (contig: NZ_ADNC01000004.1)	10188-10745 (locus tag: MALL_RS04280) (contig: NZ_ADNC01000022.1)
M. alvi ATCC 29626	NZ_JNJU00000000.1*	complement (52254-53027) (locus tag: T383_RS0101300) (contig: NZ_JNJU01000002.1)		complement (50239-50694) (locus tag: T383_RS0101285) (contig: NZ_JNJU01000002.1)	138685-139275 (locus tag: T383_RS0100510) (contig: NZ_JNJU01000001.1)
M. anatis 1340	NZ_AFVJ00000000.1*	7192-7998 (locus tag: GIG_RS00500) (contig: NZ_AFVJ01000007.1)	6078-7205 (locus tag: GIG_RS00495) (contig: NZ_AFVJ01000007.1)	21008-21439 (locus tag: GIG_RS03275) (contig: NZ_AFVJ01000035.1)	complement (15831-16403) (locus tag: GIG_RS04180) (contig: NZ_AFVJ01000038.1)
M. arginini ATCC 23838	NZ_AUAH00000000.1*	28237-28989 (locus tag: F805_RS0103520) (contig: NZ_KE386772.1)	27107-28261 (locus tag: F805_RS0103515) (contig: NZ_KE386772.1)	18799-19230 (locus tag: F805_RS0101715) (contig: NZ_AUAH01000004.1)	complement (65779-66285) (locus tag: F805_RS0100895) (contig: NZ_AUAH01000002.1)
M. bovigenitalium 51080	NZ_AP017902.1^			133462-133884 (locus tag: MBVG596_RS00500)	complement (403329-403901) (locus tag: MBVG596_RS01755)
M. bovis 08M	NZ_CP019639.1^			176421-176843 (locus tag: B0W43_RS00780)	complement (910577-911149) (locus tag: B0W43_RS03820)
M. bovoculi M165/69	NZ_CP007154.1^				467626-468960 (locus tag: MYB_RS01885)
M. buteonis ATCC 51371	NZ_JPOK00000000.1*	complement (27614-28354) (locus tag: EI91_RS01795) (contig: NZ_JPOK01000006.1)		complement (32640-33086) (locus tag: EI91_RS02235) (contig: NZ_JPOK01000007.1)	complement (146386-146955) (locus tag: EI91_RS00590) (contig: NZ_JPOK01000004.1)
M. californicum ST-6	NZ_CP007521.1^			complement (380505-380936) (locus tag: MCFN_RS01545)	complement (81250-81822) (locus tag: MCFN_RS00370)
M. canis UF33	NZ_AJFS00000000.1*			complement (30033-30464) (locus tag: MCANUF33_RS00905) (contig: NZ_AJFS01000003.1)	complement (270341-270910) (locus tag: MCANUF33_RS02435) (contig: NZ_AJFS01000004.1)
M. capricolum subsp. capricolum ATCC 27343	NC_007633.1^			281447-281869 (locus tag: MCAP_RS01170)	68116-68670 (locus tag: MCAP_RS00290)
M. capricolum subsp. capripneumoniae	NZ_CP019061.1^			309851-310272 (locus tag: BVA24_RS01255)	83588-84141 (locus tag: BVA24_RS00310)

***Mycoplasma* species**	**NCBI accession number**	**Genomic location**			
		PanK type III	**CoaBC**	**PPAT**	**DPCK**
M. collis ATCC 35278	NZ_JNJY00000000.1*			13860-14303 (locus tag: T403_RS0103270) (contig: NZ_JNJY01000016.1)	complement (93474-94022) (locus tag: T403_RS0100515) (contig: NZ_JNJY01000001.1)
M. columbinum ATCC 29257	NZ_JONY00000000.1*			complement (198287-198715) (locus tag: T406_RS0100840) (contig: NZ_JONY01000003.1)	295043-295615 (locus tag: T406_RS0101290) (contig: NZ_JONY01000003.1)
M. columborale ATCC 29258	NZ_JNJZ00000000.1*	complement (1358-2179) (locus tag: T404_RS02275) (contig: NZ_JNJZ01000008.1)	complement (2167-3300) (locus tag: T404_RS02280) (contig: NZ_JNJZ01000008.1)	44390-44818 (locus tag: T404_RS02420) (contig: NZ_JNJZ01000008.1)	78644-79201 (locus tag: T404_RS01040) (contig: NZ_JNJZ01000004.1)
M. conjunctivae HRC/581T	NC_012806.1^			90586-91032 (locus tag: MCJ_RS00410)	697230-698567 (locus tag: MCJ_RS02820)
M. cricetuli ATCC 35279	NZ_JAHB00000000.1*	complement (56507-57325) (locus tag: U744_RS0100350) (contig: NZ_JAHB01000001.1)	complement (57313-58446) (locus tag: U744_RS0100355) (contig: NZ_JAHB01000001.1)	complement (45679-46107) (locus tag: U744_RS0100285) (contig: NZ_JAHB01000001.1)	complement (32973-33548) (locus tag: U744_RS0102240) (contig: NZ_JAHB01000006.1)
M. crocodyli MP145	NC_014014.1^	complement (29628-30404) (locus tag: MCRO_RS00130)		complement (624066-624506) (locus tag: MCRO_RS02700)	658005-658577 (locus tag: MCRO_RS02845)
M. dispar ATCC 27140	NZ_CP007229.1^				complement (504206-505549) (locus tag: MDIS_RS01920)
M. felifaucium ATCC 43428	NZ_JHXS00000000.1*			complement (20875-21315) (locus tag: T390_RS0102395) (contig: NZ_JHXS01000011.1)	4781-5353 (locus tag: T390_RS0100040) (contig: NZ_JHXS01000001.1)
M. felis ATCC 23391	NZ_JNKA00000000.1*			complement (7696-8109) (locus tag: T405_RS01240) (contig: NZ_JNKA01000007.1)	complement (22907-23461) (locus tag: T405_RS01860) (contig: NZ_JNKA01000013.1)
M. fermentans PG18	NC_021002.1^			complement (240477-240905) (locus tag: MBIO_RS01255)	136897-137469 (locus tag: MBIO_RS00730)
M. flocculare ATCC 27399	NZ_CP007585.1^				complement (336307-337650) (locus tag: MYF_RS01250)
M. gallinaceum B2096 8B	CP011021.1^	755465-756211 (locus tag: VO56_02800)	754326-755483 (locus tag: VO56_02795)	585022-585459 (locus tag: VO56_02005)	
M. gallinarum Mgn_IPT	NZ_LVLH00000000.1*			complement (52410-52835) (locus tag: MGALLINA_RS01585) (contig: NZ_LVLH01000028.1)	28062-28631 (locus tag: MGALLINA_RS01505) (contig: NZ_LVLH01000028.1)
M. gallisepticum str. R(low)	NC_004829.2^				346439-347041 (locus tag: MGA_RS01410)
M. genitalium G37	NC_000908.2^				321046-321642 (locus tag: MG_RS01570)

***Mycoplasma* species**	**NCBI accession number**	**Genomic location**				
		PanK type III	**CoaBC**		**PPAT**	**DPCK**
M. hyopneumoniae J	NC_007295.1^					complement (420694-422034) (locus tag: MHJ_RS01875)
M. hyorhinis HUB-1	NC_014448.1^					775007-775594 (locus tag: MHR_RS03640)
M. imitans ATCC 51306	NZ_JADI00000000.1*					16155-16745 (locus tag: P690_RS0100995) (contig: NZ_JADI01000008.1)
M. iners ATCC 19705	NZ_JNJW00000000.1*				complement (3298-3741) (locus tag: T395_RS0102600) (contig: NZ_JNJW01000012.1)	complement (13826-14398) (locus tag: T395_RS0102195) (contig: NZ_JNJW01000007.1)
M. iowae 695	NZ_AGFP00000000.1*	654-1430 (locus tag: GUU_RS04285) (contig: NZ_AGFP01000063.1)	**PPCS**: complement (2770-3480) (locus tag: GUU_RS04050) (contig: NZ_AGFP01000058.1)	**PPCDC**: complement (2229-2765) (locus tag: GUU_RS04045) (contig: NZ_AGFP01000058.1)	complement (64103-64537) (locus tag: GUU_RS00980) (contig: NZ_AGFP01000008.1)	complement (17353-17964) (locus tag: GUU_01127) (contig: NZ_AGFP01000008.1)
M. leachii PG50	NC_014751.1^				329605-330027 (locus tag: MSB_RS01360)	109349-109906 (locus tag: MSB_RS00465)
M. leonicaptivi ATCC 49890	NZ_JHWE00000000.1*				2568-2981 (locus tag: BU00_RS0101800) (contig: NZ_JHWE01000024.1)	complement (7331-7894) (locus tag: BU00_RS04890) (contig: NZ_JHWE01000024.1)
M. lipofaciens ATCC 35015	NZ_JMKY00000000.1*				complement (123928-124350) (locus tag: T388_RS0100600) (contig: NZ_JMKY01000001.1)	66707-67279 (locus tag: T388_RS0102550) (contig: NZ_JMKY01000006.1)
M. mobile 163K	NC_006908.1^	complement (725838-726569) (locus tag: MMOB_RS03205)	complement (726554-727636) (locus tag: MMOB_RS03210)		712814-713251 (locus tag: MMOB_RS03530)	713315-713878 (locus tag: MMOB_RS03135)
M. molare ATCC 27746	NZ_JHWG00000000.1*	80872-81600 (locus tag: BU19_RS0101285) (contig: NZ_JHWG01000002.1)			20255-20686 (locus tag: BU19_RS0102135) (contig: NZ_JHWG01000009.1)	complement (101826-102374) (locus tag: BU19_RS0100565) (contig: NZ_JHWG01000001.1)
M. mycoides subsp. *capri* LC str. 95010	NC_015431.1^				complement (349967-350389) (locus tag: MLC_RS01385)	128078-128647 (locus tag: MLC_RS00500)
M. mycoides subsp. *mycoides* SC str. PG1	NC_005364.2^				312742-313164 (locus tag: MSC_0272)	114548-115117 (locus tag: MSC_0096)
M. opalescens ATCC 27921	NZ_JOOB00000000.1*				21433-21870 (locus tag: T385_RS0103170) (contig: NZ_KL544020.1)	108837-109400 (locus tag: T385_RS0102295) (contig: NZ_KL544019.1)

***Mycoplasma* species**	**NCBI accession number**	**Genomic location**			
		PanK type III	**CoaBC**	**PPAT**	**DPCK**
M. ovipneumoniae SC01	NZ_AFHO00000000.1*				complement (11215-12558) (locus tag: MOSC01_RS0103605) (contig: NZ_AFHO01000022.1)
M. penetrans HF-2	NC_004432.1^	complement (1332546-1333313) (locus tag: MYPE_RS05220)		complement (1285856-1286308) (locus tag: MYPE_RS04950)	253894-254514 (locus tag: MYPE_RS05360)
M. pirum MPI25960	NZ_AZHZ00000000.1*	complement (179214-179984) (locus tag: X558_RS0103155) (contig: NZ_KK365982.1)		complement (175096-175551) (locus tag: X558_RS0103135) (contig: NZ_KK365982.1)	171820-172404 (locus tag: X558_RS04315) (contig: NZ_KK365981.1)
M. pneumoniae M129	NC_000912.1^				458998-459600 (locus tag: MPN382)
M. primatum ATCC 25948	NZ_JNJV00000000.1*			complement (4243-4665) (locus tag: T386_RS0103800) (contig: NZ_JNJV01000038.1)	23196-23768 (locus tag: T386_RS0100120) (contig: NZ_JNJV01000001.1)
M. pulmonis UAB CTIP	NC_002771.1^	complement (570749-571423) (locus tag: MYPU_RS02370)		102808-103257 (locus tag: MYPU_RS00470)	882270-882788 (locus tag: MYPU_RS03650)
M. putrefaciens KS1	NC_015946.1^			complement (615746-616171) (locus tag: MPUT_RS02695)	503051-503611 (locus tag: MPUT_RS02170)
M. simbae ATCC 49888	NZ_JNJG00000000.1*			23328-23756 (locus tag: T329_RS0101755) (contig: NZ_JNJG01000006.1)	complement (125082-125645) (locus tag: T329_RS0101475) (contig: NZ_JNJG01000005.1)
M. sturni DSM 22021	NZ_JNIZ00000000.1*	242058-242879 (locus tag: Q349_RS01120) (contig: NZ_JNIZ01000003.1)	240940-242070 (locus tag: Q349_RS01115) (contig: NZ_JNIZ01000003.1)	complement (280184-280618) (locus tag: Q349_RS01320) (contig: NZ_JNIZ01000003.1)	complement (143690-144253) (locus tag: Q349_RS02155) (contig: NZ_KL370785.1)
M. synoviae 53	NC_007294.1^	complement (8415-9167) (locus tag: MS53_RS00030)	complement (9143-10297) (locus tag: MS53_RS00035)	complement (756186-756617) (locus tag: MS53_RS03405)	complement (622193-622699) (locus tag: MS53_RS02725)
M. testudinis ATCC 43263	NZ_JHXT00000000.1*	1820-2566 (locus tag: T384_RS0104425) (contig: NZ_JHXT01000017.1)	684-1826 (locus tag: T384_RS0104420) (contig: NZ_JHXT01000017.1)	42922-43377 (locus tag: T384_RS06450) (contig: NZ_JHXT01000013.1)	13100-13717 (locus tag: T384_RS0102290) (contig: NZ_JHXT01000007.1)
M. yeatsii GM274B	NZ_CP007520.1^			complement (512943-513365) (locus tag: MYE_RS02185)	complement (301476-302039) (locus tag: MYE_RS01275)

^Complete genome (deposited to NCBI)
*Contig assembled genome (deposited to NCBI)

Supplementary Table 1.3 The CDD search results of the PanK type III amino acid sequences of the currently annotated CoA biosynthetic pathway enzyme-encoding genes and newly identified homologues

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. alligatoris	non-specific	9	190	8.15e-17	pfam03309	Pan_kinase	cl17037	Type III pantothenate kinase
	superfamily	9	190	8.15e-17	cl17037	NBD_sugar-kinase_HSP70_actin superfamily	-	Nucleotide-Binding Domain of the sugar kinase/HSP70/actin superfamily
	non-specific	9	249	1.80e-15	PRK13320	PRK13320	cl17037	pantothenate kinase
	non-specific	8	190	5.16e-11	COG1521	CoaX	cl17037	Pantothenate kinase type III
	non-specific	7	253	3.09e-10	PRK13321	PRK13321	cl17037	pantothenate kinase
	non-specific	8	190	7.94e-10	TIGR00671	baf	cl17037	pantothenate kinase, type III
	non-specific	9	190	2.26e-05	PRK13318	PRK13318	cl17037	pantothenate kinase
	non-specific	8	229	4.00e-05	PRK13324	PRK13324	cl17037	pantothenate kinase
M. alvi	non-specific	9	197	2.86e-26	pfam03309	Pan_kinase	cl17037	Type III pantothenate kinase
	superfamily	9	197	2.86e-26	cl17037	NBD_sugar-kinase_HSP70_actin superfamily	-	Nucleotide-Binding Domain of the sugar kinase/HSP70/actin superfamily
	non-specific	9	247	1.01e-17	PRK13320	PRK13320	cl17037	pantothenate kinase
	non-specific	9	255	1.05e-17	COG1521	CoaX	cl17037	Pantothenate kinase type III
	non-specific	8	196	7.61e-17	PRK13321	PRK13321	cl17037	pantothenate kinase
	non-specific	9	255	6.31e-13	PRK13318	PRK13318	cl17037	pantothenate kinase
	non-specific	9	225	8.06e-12	TIGR00671	baf	cl17037	pantothenate kinase, type III
	non-specific	8	256	3.44e-07	PRK13324	PRK13324	cl17037	pantothenate kinase
	specific	9	156	5.96e-05	cd00012	NBD_sugar-kinase_HSP70_actin	cl17037	Nucleotide-Binding Domain of the sugar kinase/HSP70/actin superfamily
	non-specific	1	225	0.000472	PRK13326	PRK13326	cl17037	pantothenate kinase
	non-specific	1	247	0.00078	PRK13331	PRK13331	cl17037	pantothenate kinase
M. anatis	-	-	-	-	-	-	-	-
M. arginini	non-specific	10	249	8.45e-22	COG1521	CoaX	cl17037	Pantothenate kinase type III
	superfamily	10	249	8.45e-22	cl17037	NBD_sugar-kinase_HSP70_actin superfamily	-	Nucleotide-Binding Domain of the sugar kinase/HSP70/actin superfamily
	non-specific	10	205	9.42e-21	pfam03309	Pan_kinase	cl17037	Type III pantothenate kinase
	non-specific	10	246	3.67e-17	PRK13320	PRK13320	cl17037	pantothenate kinase
	non-specific	10	241	5.97e-14	TIGR00671	baf	cl17037	pantothenate kinase, type III
	non-specific	10	250	1.08e-11	PRK13318	PRK13318	cl17037	pantothenate kinase
	non-specific	8	239	4.15e-06	PRK13326	PRK13326	cl17037	pantothenate kinase
	non-specific	10	249	0.000186	PRK13324	PRK13324	cl17037	pantothenate kinase
M. buteonis	non-specific	5	198	1.60e-06	pfam03309	Pan_kinase	cl17037	Type III pantothenate kinase
	superfamily	5	198	1.60e-06	cl17037	NBD_sugar-kinase_HSP70_actin superfamily	-	Nucleotide-Binding Domain of the sugar kinase/HSP70/actin superfamily
	non-specific	5	234	3.26e-06	PRK13320	PRK13320	cl17037	pantothenate kinase
	non-specific	5	235	6.44e-06	TIGR00671	baf	cl17037	pantothenate kinase, type III
	non-specific	5	234	8.74e-06	COG1521	CoaX	cl17037	Pantothenate kinase type III

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. columborale	non-specific	6	209	1.18e-12	pfam03309	Pan_kinase	cl17037	Type III pantothenate kinase
	superfamily	6	209	1.18e-12	cl17037	NBD_sugar-kinase_HSP70_actin superfamily	-	Nucleotide-Binding Domain of the sugar kinase/HSP70/actin superfamily
	non-specific	3	232	5.53e-10	COG1521	CoaX	cl17037	Pantothenate kinase type III
	non-specific	4	233	1.26e-06	TIGR00671	baf	cl17037	pantothenate kinase, type III
	non-specific	1	235	0.000122	PRK13320	PRK13320	cl17037	pantothenate kinase
	non-specific	4	223	0.00372	PRK13321	PRK13321	cl17037	pantothenate kinase
M. cricetuli	non-specific	4	233	5.58e-10	COG1521	CoaX	cl17037	Pantothenate kinase type III
	superfamily	4	233	5.58e-10	cl17037	NBD_sugar-kinase_HSP70_actin superfamily	-	Nucleotide-Binding Domain of the sugar kinase/HSP70/actin superfamily
	non-specific	112	234	1.08e-06	TIGR00671	baf	cl17037	pantothenate kinase, type III
	non-specific	5	209	9.25e-05	pfam03309	Pan_kinase	cl17037	Type III pantothenate kinase
	non-specific	189	249	0.004086	PRK13320	PRK13320	cl17037	pantothenate kinase
M. crocodyli	non-specific	14	195	3.44e-15	pfam03309	Pan_kinase	cl17037	Type III pantothenate kinase
	superfamily	14	195	3.44e-15	cl17037	NBD_sugar-kinase_HSP70_actin superfamily	-	Nucleotide-Binding Domain of the sugar kinase/HSP70/actin superfamily
	non-specific	14	253	1.86e-13	TIGR00671	baf	cl17037	pantothenate kinase, type III
	non-specific	13	256	1.91e-13	COG1521	CoaX	cl17037	Pantothenate kinase type III
	non-specific	12	257	6.11e-13	PRK13320	PRK13320	cl17037	pantothenate kinase
	non-specific	14	253	1.58e-07	PRK13321	PRK13321	cl17037	pantothenate kinase
	non-specific	14	248	5.79e-06	PRK13324	PRK13324	cl17037	pantothenate kinase
M. gallinaceum	non-specific	9	195	5.46e-16	pfam03309	Pan_kinase	cl17037	Type III pantothenate kinase
	superfamily	9	195	5.46e-16	cl17037	NBD_sugar-kinase_HSP70_actin superfamily	-	Nucleotide-Binding Domain of the sugar kinase/HSP70/actin superfamily
	non-specific	6	245	3.79e-15	COG1521	CoaX	cl17037	Pantothenate kinase type III
	non-specific	6	245	1.22e-09	PRK13320	PRK13320	cl17037	pantothenate kinase
	non-specific	9	196	7.81e-08	TIGR00671	baf	cl17037	pantothenate kinase, type III
	non-specific	6	248	4.54e-07	PRK13318	PRK13318	cl17037	pantothenate kinase
	non-specific	7	245	3.06e-06	PRK13321	PRK13321	cl17037	pantothenate kinase
	non-specific	6	212	0.000497	PRK13324	PRK13324	cl17037	pantothenate kinase
M. iowae	non-specific	6	204	3.30e-26	pfam03309	Pan_kinase	cl17037	Type III pantothenate kinase
	superfamily	6	204	3.30e-26	cl17037	NBD_sugar-kinase_HSP70_actin superfamily	-	Nucleotide-Binding Domain of the sugar kinase/HSP70/actin superfamily
	non-specific	4	251	1.33e-20	COG1521	CoaX	cl17037	Pantothenate kinase type III
	non-specific	5	191	3.98e-17	PRK13318	PRK13318	cl17037	pantothenate kinase
	non-specific	6	249	3.70e-15	TIGR00671	baf	cl17037	pantothenate kinase, type III
	non-specific	2	252	2.04e-13	PRK13320	PRK13320	cl17037	pantothenate kinase
	non-specific	5	198	2.95e-08	PRK13324	PRK13324	cl17037	pantothenate kinase
	non-specific	5	191	8.10e-08	PRK13321	PRK13321	cl17037	pantothenate kinase
M. mobile	specific	3	197	4.16e-45	pfam03309	Pan_kinase	cl17037	Type III pantothenate kinase
	superfamily	3	197	4.16e-45	cl17037	NBD_sugar-kinase_HSP70_actin superfamily	-	Nucleotide-Binding Domain of the sugar kinase/HSP70/actin superfamily
	non-specific	1	235	3.89e-38	COG1521	CoaX	cl17037	Pantothenate kinase type III

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. mobile	non-specific	1	240	1.19e-27	PRK13318	PRK13318	cl17037	pantothenate kinase
	non-specific	1	240	6.36e-26	PRK13321	PRK13321	cl17037	pantothenate kinase
	non-specific	3	236	7.86e-25	TIGR00671	baf	cl17037	pantothenate kinase, type III
	non-specific	1	235	1.87e-14	PRK13320	PRK13320	cl17037	pantothenate kinase
	non-specific	3	217	3.77e-11	PRK13326	PRK13326	cl17037	pantothenate kinase
	non-specific	1	219	5.53e-10	PRK13324	PRK13324	cl17037	pantothenate kinase
	non-specific	3	236	1.08e-07	PRK13331	PRK13331	cl17037	pantothenate kinase
	non-specific	118	241	2.91e-07	PRK13322	PRK13322	cl17037	pantothenate kinase
	non-specific	93	235	1.66e-05	PRK13328	PRK13328	cl17037	pantothenate kinase
	non-specific	6	236	0.000303	PRK13333	PRK13333	cl17037	pantothenate kinase
	non-specific	31	124	0.007289	cd14228	STKc_HIPK1	cl21453	Catalytic domain of the Serine/Threonine Kinase, Homeodomain-Interacting Protein Kinase 1
	superfamily	31	124	0.007289	cl21453	PKc_like superfamily	-	Protein Kinases, catalytic domain
M. molare	non-specific	1	231	3.20e-23	COG1521	CoaX	cl17037	Pantothenate kinase type III
	superfamily	1	231	3.20e-23	cl17037	NBD_sugar-kinase_HSP70_actin superfamily	-	Nucleotide-Binding Domain of the sugar kinase/HSP70/actin superfamily
	non-specific	2	198	2.50e-19	pfam03309	Pan_kinase	cl17037	Type III pantothenate kinase
	non-specific	1	228	7.15e-18	TIGR00671	baf	cl17037	pantothenate kinase, type III
	non-specific	37	213	1.39e-15	PRK13318	PRK13318	cl17037	pantothenate kinase
	non-specific	116	231	6.27e-10	PRK13333	PRK13333	cl17037	pantothenate kinase
	non-specific	2	212	8.31e-08	PRK13321	PRK13321	cl17037	pantothenate kinase
	non-specific	1	229	1.20e-05	PRK13324	PRK13324	cl17037	pantothenate kinase
	non-specific	2	222	1.45e-05	PRK13320	PRK13320	cl17037	pantothenate kinase
	non-specific	2	226	0.00397	PRK13326	PRK13326	cl17037	pantothenate kinase
M. penetrans	specific	9	193	1.56e-33	pfam03309	Pan_kinase	cl17037	Type III pantothenate kinase
	superfamily	9	193	1.56e-33	cl17037	NBD_sugar-kinase_HSP70_actin superfamily	-	Nucleotide-Binding Domain of the sugar kinase/HSP70/actin superfamily
	non-specific	8	254	4.34e-21	COG1521	CoaX	cl17037	Pantothenate kinase type III
	non-specific	8	255	2.72e-16	PRK13320	PRK13320	cl17037	pantothenate kinase
	non-specific	9	226	3.94e-16	TIGR00671	baf	cl17037	pantothenate kinase, type III
	non-specific	8	224	2.33e-15	PRK13318	PRK13318	cl17037	pantothenate kinase
	non-specific	7	254	3.06e-15	PRK13321	PRK13321	cl17037	pantothenate kinase
	non-specific	8	255	1.80e-11	PRK13324	PRK13324	cl17037	pantothenate kinase
	non-specific	9	227	0.00134	PRK13326	PRK13326	cl17037	pantothenate kinase
	non-specific	126	193	0.001658	PRK13331	PRK13331	cl17037	pantothenate kinase
M. pirum	non-specific	9	193	1.71e-16	pfam03309	Pan_kinase	cl17037	Type III pantothenate kinase
	superfamily	9	193	1.71e-16	cl17037	NBD_sugar-kinase_HSP70_actin superfamily	-	Nucleotide-Binding Domain of the sugar kinase/HSP70/actin superfamily
	non-specific	9	217	5.01e-13	COG1521	CoaX	cl17037	Pantothenate kinase type III
	non-specific	9	227	1.27e-12	TIGR00671	baf	cl17037	pantothenate kinase, type III
	non-specific	9	253	3.76e-12	PRK13320	PRK13320	cl17037	pantothenate kinase
	non-specific	8	193	1.85e-06	PRK13321	PRK13321	cl17037	pantothenate kinase
	non-specific	9	193	1.12e-05	PRK13318	PRK13318	cl17037	pantothenate kinase

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. pulmonis	specific	1	224	7.06e-60	COG1521	CoaX	cl17037	Pantothenate kinase type III
	superfamily	1	224	7.06e-60	cl17037	NBD_sugar-kinase_HSP70_actin superfamily	-	Nucleotide-Binding Domain of the sugar kinase/HSP70/actin superfamily
	non-specific	5	216	1.95e-18	TIGR00671	baf	cl17037	pantothenate kinase, type III
	non-specific	3	203	1.78e-13	pfam03309	Pan_kinase	cl17037	Type III pantothenate kinase
	non-specific	1	216	1.45e-11	PRK13321	PRK13321	cl17037	pantothenate kinase
	non-specific	1	216	4.25e-10	PRK13320	PRK13320	cl17037	pantothenate kinase
	non-specific	1	224	6.07e-07	PRK13318	PRK13318	cl17037	pantothenate kinase
	non-specific	121	216	0.000398	PRK13333	PRK13333	cl17037	pantothenate kinase
M. sturni	non-specific	6	209	5.47e-09	pfam03309	Pan_kinase	cl17037	Type III pantothenate kinase
	superfamily	6	209	5.47e-09	cl17037	NBD_sugar-kinase_HSP70_actin superfamily	-	Nucleotide-Binding Domain of the sugar kinase/HSP70/actin superfamily
	non-specific	4	211	2.00e-05	PRK13320	PRK13320	cl17037	pantothenate kinase
	non-specific	6	264	2.17e-05	COG1521	CoaX	cl17037	Pantothenate kinase type III
	non-specific	6	260	7.24e-05	TIGR00671	baf	cl17037	pantothenate kinase, type III
M. synoviae	non-specific	10	249	2.74e-21	COG1521	CoaX	cl17037	Pantothenate kinase type III
	superfamily	10	249	2.74e-21	cl17037	NBD_sugar-kinase_HSP70_actin superfamily	-	Nucleotide-Binding Domain of the sugar kinase/HSP70/actin superfamily
	non-specific	10	205	3.32e-21	pfam03309	Pan_kinase	cl17037	Type III pantothenate kinase
	non-specific	10	246	1.13e-17	PRK13320	PRK13320	cl17037	pantothenate kinase
	non-specific	10	241	2.91e-14	TIGR00671	baf	cl17037	pantothenate kinase, type III
	non-specific	10	250	1.46e-11	PRK13318	PRK13318	cl17037	pantothenate kinase
	non-specific	8	239	4.27e-06	PRK13326	PRK13326	cl17037	pantothenate kinase
	non-specific	10	249	0.000354	PRK13324	PRK13324	cl17037	pantothenate kinase
M. testudinis	non-specific	5	242	1.90e-38	PRK13318	PRK13318	cl17037	pantothenate kinase
	superfamily	5	242	1.90e-38	cl17037	NBD_sugar-kinase_HSP70_actin superfamily	-	Nucleotide-Binding Domain of the sugar kinase/HSP70/actin superfamily
	non-specific	5	238	1.17e-37	PRK13321	PRK13321	cl17037	pantothenate kinase
	specific	5	201	1.90e-37	pfam03309	Pan_kinase	cl17037	Type III pantothenate kinase
	non-specific	5	242	3.20e-36	COG1521	CoaX	cl17037	Pantothenate kinase type III
	non-specific	5	241	1.68e-26	TIGR00671	baf	cl17037	pantothenate kinase, type III
	non-specific	4	242	2.53e-19	PRK13320	PRK13320	cl17037	pantothenate kinase
	non-specific	5	246	6.96e-17	PRK13326	PRK13326	cl17037	pantothenate kinase
	non-specific	5	240	8.26e-17	PRK13324	PRK13324	cl17037	pantothenate kinase
	non-specific	9	236	1.30e-12	PRK13331	PRK13331	cl17037	pantothenate kinase
	non-specific	1	240	5.88e-10	PRK13322	PRK13322	cl17037	pantothenate kinase
	non-specific	8	243	6.43e-10	PRK13333	PRK13333	cl17037	pantothenate kinase
	non-specific	100	236	0.000424	PRK13328	PRK13328	cl17037	pantothenate kinase

Supplementary Table 1.4 The CDD search results of the CoaBC amino acid sequences of the currently annotated CoA biosynthetic pathway enzyme-encoding genes and newly identified homologues

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. anatis	non-specific	2	375	6.90e-78	PRK05579	PRK05579	cl27193	bifunctional phosphopantothenoylcysteine decarboxylase/phosphopantothenate synthase
	superfamily	2	375	6.90e-78	cl27193	DFP superfamily	-	DNA / pantothenate metabolism flavoprotein
	non-specific	1	371	5.97e-71	TIGR00521	coaBC_dfp	cl27193	phosphopantothenoylcysteine decarboxylase / phosphopantothenate--cysteine ligase
	specific	1	375	5.33e-61	COG0452	CoaBC	cl27193	Phosphopantothenoylcysteine synthetase/decarboxylase
	non-specific	1	172	9.30e-37	PRK07313	PRK07313	cl19190	phosphopantothenoylcysteine decarboxylase
	superfamily	1	172	9.30e-37	cl19190	Flavoprotein superfamily	-	Flavoprotein
	non-specific	2	315	8.63e-32	PRK13982	PRK13982	cl27193	bifunctional SbtC-like/phosphopantothenoylcysteine decarboxylase/phosphopantothenate synthase
	specific	1	139	1.04e-27	pfam02441	Flavoprotein	cl19190	Flavoprotein
	non-specific	171	332	2.78e-25	pfam04127	DFP	cl27193	DNA / pantothenate metabolism flavoprotein
	non-specific	2	162	8.50e-23	TIGR02113	coaC_strep	cl19190	phosphopantothenoylcysteine decarboxylase, streptococcal
	non-specific	2	158	2.63e-21	PLN02496	PLN02496	cl19190	probable phosphopantothenoylcysteine decarboxylase
	non-specific	171	372	3.59e-09	PRK09620	PRK09620	cl27193	hypothetical protein
	non-specific	172	372	2.26e-07	PRK06732	PRK06732	cl27193	phosphopantothenate--cysteine ligase
	non-specific	172	371	2.59e-05	TIGR02114	coaB_strep	cl27193	phosphopantothenate--cysteine ligase, streptococcal
	non-specific	71	96	0.000464	PRK08305	spoVFB	cl19190	dipicolinate synthase subunit B
M. arginini	non-specific	2	373	5.24e-81	PRK05579	PRK05579	cl27193	bifunctional phosphopantothenoylcysteine decarboxylase/phosphopantothenate synthase
	superfamily	2	373	5.24e-81	cl27193	DFP superfamily	-	DNA / pantothenate metabolism flavoprotein
	specific	1	373	1.99e-68	COG0452	CoaBC	cl27193	Phosphopantothenoylcysteine synthetase/decarboxylase
	non-specific	1	340	9.94e-68	TIGR00521	coaBC_dfp	cl27193	phosphopantothenoylcysteine decarboxylase / phosphopantothenate--cysteine ligase
	non-specific	2	340	6.57e-42	PRK13982	PRK13982	cl27193	bifunctional SbtC-like/phosphopantothenoylcysteine decarboxylase/phosphopantothenate synthase
	non-specific	1	169	2.33e-36	PRK07313	PRK07313	cl19190	phosphopantothenoylcysteine decarboxylase
	superfamily	1	169	2.33e-36	cl19190	Flavoprotein superfamily	-	Flavoprotein
	specific	1	173	9.83e-35	pfam02441	Flavoprotein	cl19190	Flavoprotein
	non-specific	169	340	7.83e-30	pfam04127	DFP	cl27193	DNA / pantothenate metabolism flavoprotein
	non-specific	2	169	1.61e-25	TIGR02113	coaC_strep	cl19190	phosphopantothenoylcysteine decarboxylase, streptococcal
	non-specific	2	176	3.30e-19	PLN02496	PLN02496	cl19190	probable phosphopantothenoylcysteine decarboxylase
	non-specific	176	368	1.25e-06	TIGR02114	coaB_strep	cl27193	phosphopantothenate--cysteine ligase, streptococcal
	non-specific	176	342	3.52e-05	PRK09620	PRK09620	cl27193	hypothetical protein
	non-specific	176	368	4.15e-05	PRK06732	PRK06732	cl27193	phosphopantothenate--cysteine ligase
M. columborale	non-specific	2	374	5.64e-92	PRK05579	PRK05579	cl27193	bifunctional phosphopantothenoylcysteine decarboxylase/phosphopantothenate synthase
	superfamily	2	374	5.64e-92	cl27193	DFP superfamily	-	DNA / pantothenate metabolism flavoprotein
	non-specific	1	371	1.57e-80	TIGR00521	coaBC_dfp	cl27193	phosphopantothenoylcysteine decarboxylase / phosphopantothenate--cysteine ligase

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. columborale	specific	1	375	3.89e-74	COG0452	CoaBC	cl27193	Phosphopantothenoylcysteine synthetase/decarboxylase
	non-specific	2	334	3.07e-44	PRK13982	PRK13982	cl27193	bifunctional SbtC-like/phosphopantothenoylcysteine decarboxylase/phosphopantothenate synthase
	non-specific	1	167	1.76e-38	PRK07313	PRK07313	cl19190	phosphopantothenoylcysteine decarboxylase
	superfamily	1	167	1.76e-38	cl19190	Flavoprotein superfamily	-	Flavoprotein
	specific	1	169	1.61e-35	pfam02441	Flavoprotein	cl19190	Flavoprotein
	non-specific	171	332	2.35e-35	pfam04127	DFP	cl27193	DNA / pantothenate metabolism flavoprotein
	non-specific	2	153	3.16e-25	TIGR02113	coaC_strep	cl19190	phosphopantothenoylcysteine decarboxylase, streptococcal
	non-specific	2	168	3.40e-23	PLN02496	PLN02496	cl19190	probable phosphopantothenoylcysteine decarboxylase
	non-specific	172	263	3.28e-06	TIGR02114	coaB_strep	cl27193	phosphopantothenate--cysteine ligase, streptococcal
	non-specific	172	263	4.86e-06	PRK06732	PRK06732	cl27193	phosphopantothenate--cysteine ligase
	non-specific	1	102	3.18e-05	COG0163	UbiX	cl19190	3-polyprenyl-4-hydroxybenzoate decarboxylase
	non-specific	1	102	0.000164	PRK05920	PRK05920	cl19190	aromatic acid decarboxylase
	non-specific	2	140	0.007562	TIGR00421	ubiX_pad	cl19190	polyprenyl P-hydroxybenzoate and phenylacrylic acid decarboxylases
	non-specific	1	102	0.009011	PRK06029	PRK06029	cl19190	3-octaprenyl-4-hydroxybenzoate carboxy-lyase
M. cricetuli	non-specific	2	377	8.21e-88	PRK05579	PRK05579	cl27193	bifunctional phosphopantothenoylcysteine decarboxylase/phosphopantothenate synthase
	superfamily	2	377	8.21e-88	cl27193	DFP superfamily	-	DNA / pantothenate metabolism flavoprotein
	non-specific	1	367	2.67e-73	TIGR00521	coaBC_dfp	cl27193	phosphopantothenoylcysteine decarboxylase / phosphopantothenate--cysteine ligase
	specific	1	376	7.41e-72	COG0452	CoaBC	cl27193	Phosphopantothenoylcysteine synthetase/decarboxylase
	non-specific	2	371	2.81e-41	PRK13982	PRK13982	cl27193	bifunctional SbtC-like/phosphopantothenoylcysteine decarboxylase/phosphopantothenate synthase
	specific	169	332	1.02e-37	pfam04127	DFP	cl27193	DNA / pantothenate metabolism flavoprotein
	non-specific	1	167	1.27e-33	PRK07313	PRK07313	cl19190	phosphopantothenoylcysteine decarboxylase
	superfamily	1	167	1.27e-33	cl19190	Flavoprotein superfamily	-	Flavoprotein
	non-specific	1	168	8.96e-31	pfam02441	Flavoprotein	cl19190	Flavoprotein
	non-specific	2	158	4.89e-17	PLN02496	PLN02496	cl19190	probable phosphopantothenoylcysteine decarboxylase
	non-specific	2	167	7.40e-17	TIGR02113	coaC_strep	cl19190	phosphopantothenoylcysteine decarboxylase, streptococcal
	non-specific	169	374	3.40e-07	PRK09620	PRK09620	cl27193	hypothetical protein
	non-specific	172	372	1.44e-06	PRK06732	PRK06732	cl27193	phosphopantothenate--cysteine ligase
	non-specific	1	81	0.003177	pfam02525	Flavodoxin_2	cl00438	Flavodoxin-like fold
	superfamily	1	81	0.003177	cl00438	FMN_red superfamily	-	NADPH-dependent FMN reductase
M. gallinaceum	non-specific	2	339	2.03e-74	PRK05579	PRK05579	cl27193	bifunctional phosphopantothenoylcysteine decarboxylase/phosphopantothenate synthase
	superfamily	2	339	2.03e-74	cl27193	DFP superfamily	-	DNA / pantothenate metabolism flavoprotein
	specific	1	340	1.50e-59	COG0452	CoaBC	cl27193	Phosphopantothenoylcysteine synthetase/decarboxylase
	non-specific	1	378	1.60e-59	TIGR00521	coaBC_dfp	cl27193	phosphopantothenoylcysteine decarboxylase / phosphopantothenate--cysteine ligase
	non-specific	1	179	1.07e-38	PRK07313	PRK07313	cl19190	phosphopantothenoylcysteine decarboxylase
	superfamily	1	179	1.07e-38	cl19190	Flavoprotein superfamily	-	Flavoprotein

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. gallinaceum	non-specific	4	339	1.22e-33	PRK13982	PRK13982	cl27193	bifunctional SbtC-like/phosphopantothenoylcysteine decarboxylase/phosphopantothenate synthase
	non-specific	2	160	8.43e-32	TIGR02113	coaC_strep	cl19190	phosphopantothenoylcysteine decarboxylase, streptococcal
	specific	1	172	4.18e-26	pfam02441	Flavoprotein	cl19190	Flavoprotein
	non-specific	179	339	6.89e-26	pfam04127	DFP	cl27193	DNA / pantothenate metabolism flavoprotein
	non-specific	2	165	1.12e-17	PLN02496	PLN02496	cl19190	probable phosphopantothenoylcysteine decarboxylase
	non-specific	179	377	4.29e-06	PRK06732	PRK06732	cl27193	phosphopantothenate--cysteine ligase
	non-specific	1	103	0.000137	COG0163	UbiX	cl19190	3-polyprenyl-4-hydroxybenzoate decarboxylase
	non-specific	179	381	0.000214	TIGR02114	coaB_strep	cl27193	phosphopantothenate--cysteine ligase, streptococcal
	non-specific	77	109	0.000224	COG1036	COG1036	cl27425	Archaeal flavoprotein
	superfamily	77	109	0.000224	cl27425	COG1036 superfamily	-	Archaeal flavoprotein
	non-specific	179	356	0.000777	PRK09620	PRK09620	cl27193	hypothetical protein
	non-specific	76	107	0.001221	TIGR02700	flavo_MJ0208	cl25361	archaeoflavoprotein, MJ0208 family
	superfamily	76	107	0.001221	cl25361	NuoI superfamily	-	Formate hydrogenlyase subunit 6/NADH:ubiquinone oxidoreductase 23 kD subunit (chain I)
	non-specific	77	114	0.002808	TIGR02699	archaeo_AfpA	cl27425	archaeoflavoprotein AfpA
M. mobile	non-specific	1	356	7.51e-80	PRK05579	PRK05579	cl27193	bifunctional phosphopantothenoylcysteine decarboxylase/phosphopantothenate synthase
	superfamily	1	356	7.51e-80	cl27193	DFP superfamily	-	DNA / pantothenate metabolism flavoprotein
	non-specific	2	355	8.99e-63	TIGR00521	coaBC_dfp	cl27193	phosphopantothenoylcysteine decarboxylase / phosphopantothenate--cysteine ligase
	specific	1	356	8.69e-61	COG0452	CoaBC	cl27193	Phosphopantothenoylcysteine synthetase/decarboxylase
	non-specific	2	317	1.03e-45	PRK13982	PRK13982	cl27193	bifunctional SbtC-like/phosphopantothenoylcysteine decarboxylase/phosphopantothenate synthase
	non-specific	1	173	1.98e-32	PRK07313	PRK07313	cl19190	phosphopantothenoylcysteine decarboxylase
	superfamily	1	173	1.98e-32	cl19190	Flavoprotein superfamily	-	Flavoprotein
	specific	2	164	5.41e-32	pfam02441	Flavoprotein	cl19190	Flavoprotein
	non-specific	176	318	2.58e-26	pfam04127	DFP	cl27193	DNA / pantothenate metabolism flavoprotein
	non-specific	2	169	5.12e-21	TIGR02113	coaC_strep	cl19190	phosphopantothenoylcysteine decarboxylase, streptococcal
	non-specific	64	161	5.43e-15	PLN02496	PLN02496	cl19190	probable phosphopantothenoylcysteine decarboxylase
	non-specific	177	356	8.29e-10	PRK06732	PRK06732	cl27193	phosphopantothenate--cysteine ligase
	non-specific	177	356	9.90e-10	TIGR02114	coaB_strep	cl27193	phosphopantothenate--cysteine ligase, streptococcal
	non-specific	22	115	1.38e-05	TIGR02852	spore_dpaB	cl19190	dipicolinic acid synthetase, B subunit
	non-specific	2	115	0.000155	PRK08305	spoVFB	cl19190	dipicolinate synthase subunit B
	non-specific	70	112	0.002594	COG1036	COG1036	cl27425	Archaeal flavoprotein
	superfamily	70	112	0.002594	cl27425	COG1036 superfamily	-	Archaeal flavoprotein
	non-specific	1	95	0.008873	PRK06029	PRK06029	cl19190	3-octaprenyl-4-hydroxybenzoate carboxy-lyase
M. sturni	non-specific	2	376	1.13e-104	PRK05579	PRK05579	cl27193	bifunctional phosphopantothenoylcysteine decarboxylase/phosphopantothenate synthase
	superfamily	2	376	1.13e-104	cl27193	DFP superfamily	-	DNA / pantothenate metabolism flavoprotein
	non-specific	1	370	7.06e-81	TIGR00521	coaBC_dfp	cl27193	phosphopantothenoylcysteine decarboxylase / phosphopantothenate--cysteine ligase

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. sturni	specific	1	375	1.29e-74	COG0452	CoaBC	cl27193	Phosphopantothenoylcysteine synthetase/decarboxylase
	non-specific	2	372	2.43e-49	PRK13982	PRK13982	cl27193	bifunctional SbtC-like/phosphopantothenoylcysteine decarboxylase/phosphopantothenate synthase
	non-specific	1	165	6.92e-40	PRK07313	PRK07313	cl19190	phosphopantothenoylcysteine decarboxylase
	superfamily	1	165	6.92e-40	cl19190	Flavoprotein superfamily	-	Flavoprotein
	specific	1	158	6.94e-37	pfam02441	Flavoprotein	cl19190	Flavoprotein
	non-specific	171	346	1.82e-36	pfam04127	DFP	cl27193	DNA / pantothenate metabolism flavoprotein
	non-specific	2	153	4.43e-28	TIGR02113	coaC_strep	cl19190	phosphopantothenoylcysteine decarboxylase, streptococcal
	non-specific	3	161	1.51e-22	PLN02496	PLN02496	cl19190	probable phosphopantothenoylcysteine decarboxylase
	non-specific	172	373	5.67e-09	PRK06732	PRK06732	cl27193	phosphopantothenate--cysteine ligase
	non-specific	171	376	2.66e-07	PRK09620	PRK09620	cl27193	hypothetical protein
	non-specific	172	371	1.08e-05	TIGR02114	coaB_strep	cl27193	phosphopantothenate--cysteine ligase, streptococcal
	non-specific	17	111	0.000138	TIGR02700	flavo_MJ0208	cl25361	archaeoflavoprotein, MJ0208 family
	superfamily	17	111	0.000138	cl25361	NuoI superfamily	-	Formate hydrogenlyase subunit 6/NADH:ubiquinone oxidoreductase 23 kD subunit (chain I)
	non-specific	71	103	0.001128	COG1036	COG1036	cl27425	Archaeal flavoprotein
	superfamily	71	103	0.001128	cl27425	COG1036 superfamily	-	Archaeal flavoprotein
M. synoviae	non-specific	2	373	5.66e-80	PRK05579	PRK05579	cl27193	bifunctional phosphopantothenoylcysteine decarboxylase/phosphopantothenate synthase
	superfamily	2	373	5.66e-80	cl27193	DFP superfamily	-	DNA / pantothenate metabolism flavoprotein
	specific	1	373	7.79e-67	COG0452	CoaBC	cl27193	Phosphopantothenoylcysteine synthetase/decarboxylase
	non-specific	1	340	2.68e-66	TIGR00521	coaBC_dfp	cl27193	phosphopantothenoylcysteine decarboxylase / phosphopantothenate--cysteine ligase
	non-specific	2	340	5.08e-41	PRK13982	PRK13982	cl27193	bifunctional SbtC-like/phosphopantothenoylcysteine decarboxylase/phosphopantothenate synthase
	non-specific	1	169	5.76e-36	PRK07313	PRK07313	cl19190	phosphopantothenoylcysteine decarboxylase
	superfamily	1	169	5.76e-36	cl19190	Flavoprotein superfamily	-	Flavoprotein
	specific	1	173	3.14e-34	pfam02441	Flavoprotein	cl19190	Flavoprotein
	non-specific	169	340	8.87e-30	pfam04127	DFP	cl27193	DNA / pantothenate metabolism flavoprotein
	non-specific	2	169	6.90e-26	TIGR02113	coaC_strep	cl19190	phosphopantothenoylcysteine decarboxylase, streptococcal
	non-specific	2	176	1.45e-19	PLN02496	PLN02496	cl19190	probable phosphopantothenoylcysteine decarboxylase
	non-specific	176	368	2.41e-07	TIGR02114	coaB_strep	cl27193	phosphopantothenate--cysteine ligase, streptococcal
	non-specific	176	368	1.73e-05	PRK06732	PRK06732	cl27193	phosphopantothenate--cysteine ligase
	non-specific	176	342	6.05e-05	PRK09620	PRK09620	cl27193	hypothetical protein
M. testudinis	non-specific	17	379	2.14e-53	PRK05579	PRK05579	cl27193	bifunctional phosphopantothenoylcysteine decarboxylase/phosphopantothenate synthase
	superfamily	17	379	2.14e-53	cl27193	DFP superfamily	-	DNA / pantothenate metabolism flavoprotein
	non-specific	17	378	2.82e-51	TIGR00521	coaBC_dfp	cl27193	phosphopantothenoylcysteine decarboxylase / phosphopantothenate--cysteine ligase
	specific	17	377	1.07e-45	COG0452	CoaBC	cl27193	Phosphopantothenoylcysteine synthetase/decarboxylase
	non-specific	25	379	1.75e-33	PRK13982	PRK13982	cl27193	bifunctional SbtC-like/phosphopantothenoylcysteine decarboxylase/phosphopantothenate synthase
	non-specific	15	196	1.91e-28	PRK07313	PRK07313	cl19190	phosphopantothenoylcysteine decarboxylase

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. testudinis	superfamily	15	196	1.91e-28	cl19190	Flavoprotein superfamily	-	Flavoprotein
	specific	17	162	3.39e-22	pfam02441	Flavoprotein	cl19190	Flavoprotein
	non-specific	16	192	7.99e-14	TIGR02113	coaC_strep	cl19190	phosphopantothenoylcysteine decarboxylase, streptococcal
	non-specific	18	140	7.92e-09	PLN02496	PLN02496	cl19190	probable phosphopantothenoylcysteine decarboxylase
	non-specific	196	327	1.21e-06	pfam04127	DFP	cl27193	DNA / pantothenate metabolism flavoprotein
	non-specific	199	377	0.000432	TIGR02114	coaB_strep	cl27193	phosphopantothenate--cysteine ligase, streptococcal
	non-specific	94	161	0.001286	PRK08305	spoVFB	cl19190	dipicolinate synthase subunit B
	non-specific	200	377	0.003129	PRK06732	PRK06732	cl27193	phosphopantothenate--cysteine ligase
M. iowae (PPCDC)	non-specific	3	177	9.47e-70	PRK07313	PRK07313	cl19190	phosphopantothenoylcysteine decarboxylase
	superfamily	3	177	9.47e-70	cl19190	Flavoprotein superfamily	-	Flavoprotein
	non-specific	1	177	6.25e-64	PRK05579	PRK05579	cl27193	bifunctional phosphopantothenoylcysteine decarboxylase/phosphopantothenate synthase
	superfamily	1	177	6.25e-64	cl27193	DFP superfamily	-	DNA / pantothenate metabolism flavoprotein
	non-specific	3	176	3.91e-58	TIGR02113	coaC_strep	cl19190	phosphopantothenoylcysteine decarboxylase, streptococcal
	specific	1	177	1.86e-53	COG0452	CoaBC	cl27193	Phosphopantothenoylcysteine synthetase/decarboxylase
	non-specific	2	171	1.76e-51	TIGR00521	coaBC_dfp	cl27193	phosphopantothenoylcysteine decarboxylase / phosphopantothenate--cysteine ligase
	specific	3	174	2.56e-41	pfam02441	Flavoprotein	cl19190	Flavoprotein
	non-specific	3	165	6.77e-25	PRK13982	PRK13982	cl27193	bifunctional SbtC-like/phosphopantothenoylcysteine decarboxylase/phosphopantothenate synthase
	non-specific	5	166	6.90e-24	PLN02496	PLN02496	cl19190	probable phosphopantothenoylcysteine decarboxylase
	non-specific	1	174	1.40e-07	COG0163	UbiX	cl19190	3-polyprenyl-4-hydroxybenzoate decarboxylase
	non-specific	79	176	3.10e-06	COG1036	COG1036	cl27425	Archaeal flavoprotein
	superfamily	79	176	3.10e-06	cl27425	COG1036 superfamily	-	Archaeal flavoprotein
	non-specific	79	121	2.99e-05	TIGR02699	archaeo_AfpA	cl27425	archaeoflavoprotein AfpA
	non-specific	79	174	3.73e-05	TIGR00421	ubiX_pad	cl19190	polyprenyl P-hydroxybenzoate and phenylacrylic acid decarboxylases
	non-specific	2	101	0.000127	PRK08305	spoVFB	cl19190	dipicolinate synthase subunit B
	non-specific	78	149	0.00026	PRK06029	PRK06029	cl19190	3-octaprenyl-4-hydroxybenzoate carboxy-lyase
	non-specific	1	115	0.000576	PRK05920	PRK05920	cl19190	aromatic acid decarboxylase
M. iowae (PPCS)	non-specific	1	235	4.40e-46	PRK06732	PRK06732	cl27193	phosphopantothenate--cysteine ligase
	superfamily	1	235	4.40e-46	cl27193	DFP superfamily	-	DNA / pantothenate metabolism flavoprotein
	non-specific	2	235	2.55e-32	TIGR02114	coaB_strep	cl27193	phosphopantothenate--cysteine ligase, streptococcal
	non-specific	2	233	4.06e-16	PRK09620	PRK09620	cl27193	hypothetical protein
	non-specific	2	194	6.27e-13	pfam04127	DFP	cl27193	DNA / pantothenate metabolism flavoprotein
	non-specific	2	194	9.56e-12	PRK05579	PRK05579	cl27193	bifunctional phosphopantothenoylcysteine decarboxylase/phosphopantothenate synthase
	non-specific	1	210	5.70e-10	COG0452	CoaBC	cl27193	Phosphopantothenoylcysteine synthetase/decarboxylase
	non-specific	2	194	7.80e-10	TIGR00521	coaBC_dfp	cl27193	phosphopantothenoylcysteine decarboxylase / phosphopantothenate--cysteine ligase
	non-specific	1	112	0.005721	COG0451	WcaG	cl25660	Nucleoside-diphosphate-sugar epimerase
	superfamily	1	112	0.005721	cl25660	Epimerase superfamily	-	NAD dependent epimerase/dehydratase family

Supplementary Table 1.5 The CDD search results of the PPAT amino acid sequences of the currently annotated CoA biosynthetic pathway enzyme-encoding genes and newly identified homologues

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. agalactiae	specific	1	140	7.72e-76	PRK13964	coaD	cl00015	phosphopantetheine adenylyltransferase
	superfamily	1	140	7.72e-76	cl00015	nt_trans superfamily	-	nucleotidyl transferase superfamily
	non-specific	1	134	1.14e-44	COG0669	CoaD	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	4	133	1.13e-40	cd02163	PPAT	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	4	133	5.88e-39	TIGR01510	coaD_prev_kdtB	cl00015	pantetheine-phosphate adenylyltransferase
	non-specific	1	133	6.60e-38	PRK00168	coaD	cl00015	phosphopantetheine adenylyltransferase
	specific	5	135	8.39e-18	pfam01467	CTP_transf_like	cl00015	Cytidylyltransferase-like
	non-specific	4	133	1.35e-09	cd02039	cytidylyltransferase_like	cl00015	Cytidylyltransferase-like domain
	non-specific	4	60	1.67e-07	TIGR00125	cyt_tran_rel	cl00015	cytidyltransferase-like domain
	non-specific	4	66	1.13e-06	cd02165	NMNAT	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	1	51	2.22e-06	PRK00071	nadD	cl00015	nicotinic acid mononucleotide adenylyltransferase
	non-specific	1	66	1.31e-05	PRK07152	nadD	cl28367	putative nicotinate-nucleotide adenylyltransferase
	superfamily	1	66	1.31e-05	cl28367	nadD superfamily	-	putative nicotinate-nucleotide adenylyltransferase
	non-specific	4	60	2.15e-05	cd02156	nt_trans	cl00015	nucleotidyl transferase superfamily
	non-specific	5	51	2.74e-05	TIGR00482	TIGR00482	cl00015	nicotinate (nicotinamide) nucleotide adenylyltransferase
	non-specific	1	87	4.49e-05	cd02170	cytidylyltransferase	cl00015	cytidylyltransferase
	non-specific	1	41	4.94e-05	COG1057	NadD	cl00015	Nicotinic acid mononucleotide adenylyltransferase
	non-specific	1	59	0.000196	COG1056	NadR	cl00015	Nicotinamide mononucleotide adenylyltransferase
	non-specific	5	73	0.00042	cd02167	NMNAT_NadR	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	14	39	0.000487	pfam02569	Pantoate_ligase	cl00015	Pantoate-beta-alanine ligase
	non-specific	1	68	0.002358	cd02171	G3P_Cytidylyltransferase	cl00015	glycerol-3-phosphate cytidylyltransferase
	non-specific	13	39	0.002895	PRK00380	panC	cl00015	pantoate--beta-alanine ligase
	non-specific	8	36	0.005163	COG0615	TagD	cl00015	Glycerol-3-phosphate cytidylyltransferase, cytidylyltransferase family
	non-specific	5	42	0.00664	cd02173	ECT	cl00015	CTP:phosphoethanolamine cytidylyltransferase (ECT)
	non-specific	14	67	0.006924	cd00560	PanC	cl00015	Pantoate-beta-alanine ligase
	non-specific	4	35	0.00929	PRK01153	PRK01153	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	14	39	0.009687	COG0414	PanC	cl00015	Panthothenate synthetase
M. alligatoris	non-specific	5	145	6.82e-61	PRK13964	coaD	cl00015	phosphopantetheine adenylyltransferase
	superfamily	5	145	6.82e-61	cl00015	nt_trans superfamily	-	nucleotidyl transferase superfamily
	non-specific	4	139	5.23e-43	COG0669	CoaD	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	6	138	3.36e-39	cd02163	PPAT	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	5	138	4.23e-37	PRK00168	coaD	cl00015	phosphopantetheine adenylyltransferase
	non-specific	6	138	4.91e-36	TIGR01510	coaD_prev_kdtB	cl00015	pantetheine-phosphate adenylyltransferase
	specific	8	140	4.61e-19	pfam01467	CTP_transf_like	cl00015	Cytidylyltransferase-like
	non-specific	6	138	2.11e-11	cd02039	cytidylyltransferase_like	cl00015	Cytidylyltransferase-like domain
	non-specific	6	65	1.92e-09	TIGR00125	cyt_tran_rel	cl00015	cytidyltransferase-like domain

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. alligatoris	non-specific	8	138	3.57e-08	cd02167	NMNAT_NadR	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	1	55	3.89e-08	PRK00071	nadD	cl00015	nicotinic acid mononucleotide adenylyltransferase
	non-specific	7	145	1.33e-06	cd02165	NMNAT	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	2	55	7.19e-06	COG1057	NadD	cl00015	Nicotinic acid mononucleotide adenylyltransferase
	non-specific	6	61	8.85e-06	cd02156	nt_trans	cl00015	nucleotidyl transferase superfamily
	non-specific	8	74	1.38e-05	TIGR00482	TIGR00482	cl00015	nicotinate (nicotinamide) nucleotide adenylyltransferase
	non-specific	4	71	1.89e-05	PRK07152	nadD	cl28367	putative nicotinate-nucleotide adenylyltransferase
	superfamily	4	71	1.89e-05	cl28367	nadD superfamily	-	putative nicotinate-nucleotide adenylyltransferase
	non-specific	1	139	1.92e-05	PRK08099	PRK08099	cl28365	bifunctional DNA-binding transcriptional repressor/ NMN adenylyltransferase
	superfamily	1	139	1.92e-05	cl28365	PRK08099 superfamily	-	bifunctional DNA-binding transcriptional repressor/ NMN adenylyltransferase
	non-specific	8	138	2.08e-05	TIGR01526	nadR_NMN_Atrans	cl28365	nicotinamide-nucleotide adenylyltransferase
	non-specific	3	76	4.92e-05	COG1056	NadR	cl00015	Nicotinamide mononucleotide adenylyltransferase
	non-specific	5	35	0.000102	PRK00777	PRK00777	cl00015	phosphopantetheine adenylyltransferase
	non-specific	8	45	0.000144	cd02174	CCT	cl00015	CTP:phosphocholine cytidylyltransferase
	non-specific	11	44	0.000385	TIGR02199	rfaE_dom_II	cl00015	rfaE bifunctional protein, domain II
	non-specific	5	39	0.000474	COG0615	TagD	cl00015	Glycerol-3-phosphate cytidylyltransferase, cytidylyltransferase family
	non-specific	5	39	0.000767	cd02170	cytidylyltransferase	cl00015	cytidylyltransferase
	non-specific	1	36	0.001749	COG1019	CAB4	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	6	45	0.002804	cd02173	ECT	cl00015	CTP:phosphoethanolamine cytidylyltransferase (ECT)
	non-specific	11	44	0.002804	COG2870	RfaE	cl28454	ADP-heptose synthase, bifunctional sugar kinase/adenylyltransferase
	superfamily	11	44	0.002804	cl28454	RfaE superfamily	-	ADP-heptose synthase, bifunctional sugar kinase/adenylyltransferase
	non-specific	7	38	0.003547	PRK05379	PRK05379	cl28366	bifunctional nicotinamide mononucleotide adenylyltransferase
	superfamily	7	38	0.003547	cl28366	PRK05379 superfamily	-	bifunctional nicotinamide mononucleotide adenylyltransferase
	non-specific	5	46	0.003932	cd02171	G3P_Cytidylyltransferase	cl00015	glycerol-3-phosphate cytidylyltransferase
	non-specific	11	43	0.004677	PRK11316	PRK11316	cl28454	bifunctional heptose 7-phosphate kinase/heptose 1-phosphate adenyltransferase
M. alvi	non-specific	5	136	1.05e-34	cd02163	PPAT	cl00015	Phosphopantetheine adenylyltransferase
	superfamily	5	136	1.05e-34	cl00015	nt_trans superfamily	-	nucleotidyl transferase superfamily
	non-specific	3	136	1.63e-33	COG0669	CoaD	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	3	136	1.11e-31	PRK00168	coaD	cl00015	phosphopantetheine adenylyltransferase
	non-specific	5	136	1.40e-31	TIGR01510	coaD_prev_kdtB	cl00015	pantetheine-phosphate adenylyltransferase
	non-specific	3	147	1.40e-27	PRK13964	coaD	cl00015	phosphopantetheine adenylyltransferase
	non-specific	7	137	4.55e-14	pfam01467	CTP_transf_like	cl00015	Cytidylyltransferase-like
	non-specific	5	37	3.10e-10	TIGR00125	cyt_tran_rel	cl00015	cytidyltransferase-like domain
	non-specific	1	40	2.38e-09	COG1057	NadD	cl00015	Nicotinic acid mononucleotide adenylyltransferase

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. alvi	non-specific	1	38	4.96e-09	PRK00071	nadD	cl00015	nicotinic acid mononucleotide adenylyltransferase
	non-specific	6	37	1.43e-07	cd02166	NMNAT_Archaea	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	1	37	1.53e-07	COG1056	NadR	cl00015	Nicotinamide mononucleotide adenylyltransferase
	non-specific	5	40	2.48e-07	cd02165	NMNAT	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	6	37	5.59e-07	PRK01153	PRK01153	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	3	45	1.01e-06	PRK07152	nadD	cl28367	putative nicotinate-nucleotide adenylyltransferase
	superfamily	3	45	1.01e-06	cl28367	nadD superfamily	-	putative nicotinate-nucleotide adenylyltransferase
	non-specific	6	60	3.31e-06	cd02156	nt_trans	cl00015	nucleotidyl transferase superfamily
	non-specific	5	136	4.91e-05	cd02039	cytidylyltransferase_like	cl00015	Cytidylyltransferase-like domain
	non-specific	7	40	0.000267	TIGR00482	TIGR00482	cl00015	nicotinate (nicotinamide) nucleotide adenylyltransferase
	non-specific	10	38	0.000627	COG0615	TagD	cl00015	Glycerol-3-phosphate cytidylyltransferase, cytidylyltransferase family
	non-specific	10	38	0.000673	cd02170	cytidylyltransferase	cl00015	cytidylyltransferase
	non-specific	3	38	0.000929	TIGR01526	nadR_NMN_Atrans	cl28365	nicotinamide-nucleotide adenylyltransferase
	superfamily	3	38	0.000929	cl28365	PRK08099 superfamily	-	bifunctional DNA-binding transcriptional repressor/ NMN adenylyltransferase
	non-specific	4	35	0.000937	PRK06973	PRK06973	cl00015	nicotinic acid mononucleotide adenylyltransferase
	non-specific	6	37	0.001266	TIGR01527	arch_NMN_Atrans	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	3	43	0.001379	cd02172	RfaE_N	cl00015	N-terminal domain of RfaE
	non-specific	4	37	0.003214	PTZ00308	PTZ00308	cl28626	ethanolamine-phosphate cytidylyltransferase
	superfamily	4	37	0.003214	cl28626	PLN02406 superfamily	-	ethanolamine-phosphate cytidylyltransferase
	non-specific	5	38	0.003822	cd02167	NMNAT_NadR	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	1	44	0.007022	PRK08099	PRK08099	cl28365	bifunctional DNA-binding transcriptional repressor/ NMN adenylyltransferase
M. anatis	non-specific	3	142	5.32e-63	PRK13964	coaD	cl00015	phosphopantetheine adenylyltransferase
	superfamily	3	142	5.32e-63	cl00015	nt_trans superfamily	-	nucleotidyl transferase superfamily
	non-specific	2	136	8.28e-44	COG0669	CoaD	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	6	135	1.09e-42	TIGR01510	coaD_prev_kdtB	cl00015	pantetheine-phosphate adenylyltransferase
	non-specific	6	133	4.40e-42	cd02163	PPAT	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	3	133	1.05e-39	PRK00168	coaD	cl00015	phosphopantetheine adenylyltransferase
	specific	7	137	1.87e-18	pfam01467	CTP_transf_like	cl00015	Cytidylyltransferase-like
	non-specific	6	47	4.37e-10	TIGR00125	cyt_tran_rel	cl00015	cytidyltransferase-like domain
	non-specific	3	68	1.64e-07	PRK07152	nadD	cl28367	putative nicotinate-nucleotide adenylyltransferase
	superfamily	3	68	1.64e-07	cl28367	nadD superfamily	-	putative nicotinate-nucleotide adenylyltransferase
	non-specific	5	135	2.38e-07	cd02039	cytidylyltransferase_like	cl00015	Cytidylyltransferase-like domain
	non-specific	1	57	8.37e-07	COG1057	NadD	cl00015	Nicotinic acid mononucleotide adenylyltransferase
	non-specific	6	68	1.21e-06	cd02165	NMNAT	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	1	53	2.42e-06	PRK00071	nadD	cl00015	nicotinic acid mononucleotide adenylyltransferase
	non-specific	1	37	4.98e-06	COG1056	NadR	cl00015	Nicotinamide mononucleotide adenylyltransferase

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. anatis	non-specific	6	51	4.50e-05	cd02156	nt_trans	cl00015	nucleotidyl transferase superfamily
	non-specific	6	37	4.67e-05	PRK01153	PRK01153	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	6	37	0.000133	cd02166	NMNAT_Archaea	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	7	44	0.000544	cd02173	ECT	cl00015	CTP:phosphoethanolamine cytidylyltransferase (ECT)
	non-specific	7	46	0.000563	TIGR00482	TIGR00482	cl00015	nicotinate (nicotinamide) nucleotide adenylyltransferase
	non-specific	7	135	0.001898	cd02167	NMNAT_NadR	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	6	37	0.00322	TIGR01527	arch_NMN_Atrans	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	10	88	0.003576	COG0615	TagD	cl00015	Glycerol-3-phosphate cytidylyltransferase, cytidylyltransferase family
	non-specific	10	35	0.005471	PRK00777	PRK00777	cl00015	phosphopantetheine adenylyltransferase
	non-specific	9	44	0.008923	PTZ00308	PTZ00308	cl28626	ethanolamine-phosphate cytidylyltransferase
	superfamily	9	44	0.008923	cl28626	PLN02406 superfamily	-	ethanolamine-phosphate cytidylyltransferase
M. arginini	specific	1	140	3.43e-72	PRK13964	coaD	cl00015	phosphopantetheine adenylyltransferase
	superfamily	1	140	3.43e-72	cl00015	nt_trans superfamily	-	nucleotidyl transferase superfamily
	non-specific	1	134	1.29e-46	COG0669	CoaD	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	3	133	5.90e-43	cd02163	PPAT	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	1	133	4.85e-40	PRK00168	coaD	cl00015	phosphopantetheine adenylyltransferase
	non-specific	3	133	2.98e-38	TIGR01510	coaD_prev_kdtB	cl00015	pantetheine-phosphate adenylyltransferase
	specific	5	133	3.14e-21	pfam01467	CTP_transf_like	cl00015	Cytidylyltransferase-like
	non-specific	3	35	4.10e-12	TIGR00125	cyt_tran_rel	cl00015	cytidyltransferase-like domain
	non-specific	3	133	2.72e-11	cd02039	cytidylyltransferase_like	cl00015	Cytidylyltransferase-like domain
	non-specific	1	66	6.91e-09	COG1057	NadD	cl00015	Nicotinic acid mononucleotide adenylyltransferase
	non-specific	5	133	6.44e-08	cd02167	NMNAT_NadR	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	3	39	7.00e-08	cd02156	nt_trans	cl00015	nucleotidyl transferase superfamily
	non-specific	4	66	7.56e-08	cd02165	NMNAT	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	1	41	2.06e-07	PRK00071	nadD	cl00015	nicotinic acid mononucleotide adenylyltransferase
	non-specific	1	36	3.33e-07	cd02170	cytidylyltransferase	cl00015	cytidylyltransferase
	non-specific	1	85	4.47e-07	COG0615	TagD	cl00015	Glycerol-3-phosphate cytidylyltransferase, cytidylyltransferase family
	non-specific	1	43	6.80e-07	PRK07152	nadD	cl28367	putative nicotinate-nucleotide adenylyltransferase
	superfamily	1	43	6.80e-07	cl28367	nadD superfamily	-	putative nicotinate-nucleotide adenylyltransferase
	non-specific	5	62	2.09e-06	TIGR00482	TIGR00482	cl00015	nicotinate (nicotinamide) nucleotide adenylyltransferase
	non-specific	1	35	2.18e-06	COG1056	NadR	cl00015	Nicotinamide mononucleotide adenylyltransferase
	non-specific	1	33	1.71e-05	PRK00777	PRK00777	cl00015	phosphopantetheine adenylyltransferase
	non-specific	1	36	0.000131	TIGR01526	nadR_NMN_Atrans	cl28365	nicotinamide-nucleotide adenylyltransferase
	superfamily	1	36	0.000131	cl28365	PRK08099 superfamily	-	bifunctional DNA-binding transcriptional repressor/ NMN adenylyltransferase
	non-specific	1	36	0.000394	cd02171	G3P_Cytidylyltransferase	cl00015	glycerol-3-phosphate cytidylyltransferase
	non-specific	8	41	0.000535	cd02174	CCT	cl00015	CTP:phosphocholine cytidylyltransferase

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. arginini	non-specific	3	35	0.001099	cd02166	NMNAT_Archaea	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	2	35	0.001112	PRK01153	PRK01153	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	7	41	0.00136	PTZ00308	PTZ00308	cl28626	ethanolamine-phosphate cytidylyltransferase
	superfamily	7	41	0.00136	cl28626	PLN02406 superfamily	-	ethanolamine-phosphate cytidylyltransferase
	non-specific	8	57	0.002223	TIGR01518	g3p_cytidyltrns	cl00015	glycerol-3-phosphate cytidylyltransferase
	non-specific	3	41	0.002408	cd02173	ECT	cl00015	CTP:phosphoethanolamine cytidylyltransferase (ECT)
	non-specific	8	134	0.008463	PRK08099	PRK08099	cl28365	bifunctional DNA-binding transcriptional repressor/ NMN adenylyltransferase
M. bovigenitalium	non-specific	1	140	1.28e-52	PRK13964	coaD	cl00015	phosphopantetheine adenylyltransferase
	superfamily	1	140	1.28e-52	cl00015	nt_trans superfamily	-	nucleotidyl transferase superfamily
	non-specific	1	134	1.75e-42	COG0669	CoaD	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	3	133	1.91e-40	cd02163	PPAT	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	1	133	2.71e-39	PRK00168	coaD	cl00015	phosphopantetheine adenylyltransferase
	non-specific	3	133	6.95e-37	TIGR01510	coaD_prev_kdtB	cl00015	pantetheine-phosphate adenylyltransferase
	specific	5	135	7.04e-24	pfam01467	CTP_transf_like	cl00015	Cytidylyltransferase-like
	non-specific	5	133	9.44e-14	cd02167	NMNAT_NadR	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	3	57	1.21e-12	TIGR00125	cyt_tran_rel	cl00015	cytidyltransferase-like domain
	non-specific	3	133	1.51e-10	cd02039	cytidylyltransferase_like	cl00015	Cytidylyltransferase-like domain
	non-specific	1	66	2.48e-10	PRK00071	nadD	cl00015	nicotinic acid mononucleotide adenylyltransferase
	non-specific	1	66	4.65e-10	COG1057	NadD	cl00015	Nicotinic acid mononucleotide adenylyltransferase
	non-specific	1	133	1.25e-09	TIGR01526	nadR_NMN_Atrans	cl28365	nicotinamide-nucleotide adenylyltransferase
	superfamily	1	133	1.25e-09	cl28365	PRK08099 superfamily	-	bifunctional DNA-binding transcriptional repressor/ NMN adenylyltransferase
	non-specific	4	66	9.14e-09	cd02165	NMNAT	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	1	35	3.95e-08	COG1056	NadR	cl00015	Nicotinamide mononucleotide adenylyltransferase
	non-specific	5	66	4.77e-08	TIGR00482	TIGR00482	cl00015	nicotinate (nicotinamide) nucleotide adenylyltransferase
	non-specific	1	66	7.58e-08	PRK07152	nadD	cl28367	putative nicotinate-nucleotide adenylyltransferase
	superfamily	1	66	7.58e-08	cl28367	nadD superfamily	-	putative nicotinate-nucleotide adenylyltransferase
	non-specific	1	35	3.11e-07	PRK01153	PRK01153	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	3	67	3.94e-07	cd02166	NMNAT_Archaea	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	8	134	4.62e-07	PRK08099	PRK08099	cl28365	bifunctional DNA-binding transcriptional repressor/ NMN adenylyltransferase
	non-specific	3	77	2.82e-06	TIGR01527	arch_NMN_Atrans	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	4	50	3.24e-06	PRK05379	PRK05379	cl28366	bifunctional nicotinamide mononucleotide adenylyltransferase
	superfamily	4	50	3.24e-06	cl28366	PRK05379 superfamily	-	bifunctional nicotinamide mononucleotide adenylyltransferase
	non-specific	1	33	3.44e-05	PRK00777	PRK00777	cl00015	phosphopantetheine adenylyltransferase
	non-specific	1	58	4.60e-05	COG0615	TagD	cl00015	Glycerol-3-phosphate cytidylyltransferase, cytidylyltransferase family

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. bovigenitalium	non-specific	1	56	4.75e-05	COG1019	CAB4	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	4	50	4.99e-05	cd02168	NMNAT_Nudix	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	7	56	5.04e-05	cd02164	PPAT_CoAS	cl00015	phosphopantetheine adenylyltransferase
	non-specific	3	85	0.000344	cd02156	nt_trans	cl00015	nucleotidyl transferase superfamily
	non-specific	1	133	0.000777	cd02170	cytidylyltransferase	cl00015	cytidylyltransferase
	non-specific	5	55	0.001778	cd02174	CCT	cl00015	CTP:phosphocholine cytidylyltransferase
	non-specific	1	57	0.00191	cd02171	G3P_Cytidylyltransferase	cl00015	glycerol-3-phosphate cytidylyltransferase
	non-specific	8	140	0.005664	COG0196	RibF	cl27514	FAD synthase
	superfamily	8	140	0.005664	cl27514	Flavokinase superfamily	-	Riboflavin kinase
	non-specific	8	64	0.007382	PRK07143	PRK07143	cl27514	hypothetical protein
M. bovis	specific	1	140	7.89e-71	PRK13964	coaD	cl00015	phosphopantetheine adenylyltransferase
	superfamily	1	140	7.89e-71	cl00015	nt_trans superfamily	-	nucleotidyl transferase superfamily
	non-specific	1	139	1.50e-44	COG0669	CoaD	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	4	133	6.78e-41	cd02163	PPAT	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	4	133	3.43e-38	TIGR01510	coaD_prev_kdtB	cl00015	pantetheine-phosphate adenylyltransferase
	non-specific	1	133	1.34e-37	PRK00168	coaD	cl00015	phosphopantetheine adenylyltransferase
	specific	5	135	1.14e-19	pfam01467	CTP_transf_like	cl00015	Cytidylyltransferase-like
	non-specific	4	63	2.65e-11	TIGR00125	cyt_tran_rel	cl00015	cytidyltransferase-like domain
	non-specific	4	133	3.31e-11	cd02039	cytidylyltransferase_like	cl00015	Cytidylyltransferase-like domain
	non-specific	1	97	6.97e-10	COG1056	NadR	cl00015	Nicotinamide mononucleotide adenylyltransferase
	non-specific	1	66	1.18e-08	PRK07152	nadD	cl28367	putative nicotinate-nucleotide adenylyltransferase
	superfamily	1	66	1.18e-08	cl28367	nadD superfamily	-	putative nicotinate-nucleotide adenylyltransferase
	non-specific	4	66	4.64e-08	cd02165	NMNAT	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	1	60	2.23e-07	COG1057	NadD	cl00015	Nicotinic acid mononucleotide adenylyltransferase
	non-specific	1	66	2.33e-07	PRK00071	nadD	cl00015	nicotinic acid mononucleotide adenylyltransferase
	non-specific	4	90	1.81e-06	cd02156	nt_trans	cl00015	nucleotidyl transferase superfamily
	non-specific	5	60	4.47e-06	TIGR00482	TIGR00482	cl00015	nicotinate (nicotinamide) nucleotide adenylyltransferase
	non-specific	13	39	6.07e-05	pfam02569	Pantoate_ligase	cl00015	Pantoate-beta-alanine ligase
	non-specific	5	67	6.49e-05	cd02167	NMNAT_NadR	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	4	35	0.000109	PRK01153	PRK01153	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	8	36	0.000227	cd02170	cytidylyltransferase	cl00015	cytidylyltransferase
	non-specific	4	35	0.000243	cd02166	NMNAT_Archaea	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	3	35	0.00039	PRK05379	PRK05379	cl28366	bifunctional nicotinamide mononucleotide adenylyltransferase
	superfamily	3	35	0.00039	cl28366	PRK05379 superfamily	-	bifunctional nicotinamide mononucleotide adenylyltransferase
	non-specific	8	43	0.000479	COG0615	TagD	cl00015	Glycerol-3-phosphate cytidylyltransferase, cytidylyltransferase family

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. bovis	non-specific	8	134	0.00069	PRK08099	PRK08099	cl28365	bifunctional DNA-binding transcriptional repressor/ NMN adenylyltransferase
	superfamily	8	134	0.00069	cl28365	PRK08099 superfamily	-	bifunctional DNA-binding transcriptional repressor/ NMN adenylyltransferase
	non-specific	1	68	0.001024	cd02171	G3P_Cytidylyltransferase	cl00015	glycerol-3-phosphate cytidylyltransferase
	non-specific	4	139	0.0013	TIGR01527	arch_NMN_Atrans	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	13	39	0.001386	COG0414	PanC	cl00015	Panthothenate synthetase
	non-specific	13	39	0.001429	PRK00380	panC	cl00015	pantoate--beta-alanine ligase
	non-specific	13	39	0.001477	cd00560	PanC	cl00015	Pantoate-beta-alanine ligase
	non-specific	1	49	0.001604	PRK00777	PRK00777	cl00015	phosphopantetheine adenylyltransferase
	non-specific	8	69	0.001618	COG1019	CAB4	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	1	133	0.001948	TIGR01526	nadR_NMN_Atrans	cl28365	nicotinamide-nucleotide adenylyltransferase
	non-specific	5	42	0.002332	cd02173	ECT	cl00015	CTP:phosphoethanolamine cytidylyltransferase (ECT)
	non-specific	8	52	0.002498	PRK11316	PRK11316	cl28454	bifunctional heptose 7-phosphate kinase/heptose 1-phosphate adenyltransferase
	superfamily	8	52	0.002498	cl28454	RfaE superfamily	-	ADP-heptose synthase, bifunctional sugar kinase/adenylyltransferase
	non-specific	13	39	0.002915	PLN02660	PLN02660	cl00015	pantoate--beta-alanine ligase
	non-specific	8	68	0.004536	TIGR01518	g3p_cytidyltrns	cl00015	glycerol-3-phosphate cytidylyltransferase
	non-specific	7	41	0.005859	PTZ00308	PTZ00308	cl28626	ethanolamine-phosphate cytidylyltransferase
	superfamily	7	41	0.005859	cl28626	PLN02406 superfamily	-	ethanolamine-phosphate cytidylyltransferase
M. buteonis	non-specific	5	145	1.43e-47	PRK13964	coaD	cl00015	phosphopantetheine adenylyltransferase
	superfamily	5	145	1.43e-47	cl00015	nt_trans superfamily	-	nucleotidyl transferase superfamily
	non-specific	4	141	4.02e-41	COG0669	CoaD	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	7	139	2.75e-38	cd02163	PPAT	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	7	139	7.00e-37	TIGR01510	coaD_prev_kdtB	cl00015	pantetheine-phosphate adenylyltransferase
	non-specific	5	139	1.28e-36	PRK00168	coaD	cl00015	phosphopantetheine adenylyltransferase
	specific	9	139	8.26e-18	pfam01467	CTP_transf_like	cl00015	Cytidylyltransferase-like
	non-specific	7	48	5.44e-08	TIGR00125	cyt_tran_rel	cl00015	cytidyltransferase-like domain
	non-specific	3	51	1.76e-07	COG1057	NadD	cl00015	Nicotinic acid mononucleotide adenylyltransferase
	non-specific	4	52	1.23e-06	PRK00071	nadD	cl00015	nicotinic acid mononucleotide adenylyltransferase
	non-specific	7	139	2.50e-06	cd02039	cytidylyltransferase_like	cl00015	Cytidylyltransferase-like domain
	non-specific	5	70	4.51e-06	PRK07152	nadD	cl28367	putative nicotinate-nucleotide adenylyltransferase
	superfamily	5	70	4.51e-06	cl28367	nadD superfamily	-	putative nicotinate-nucleotide adenylyltransferase
	non-specific	7	43	2.45e-05	cd02165	NMNAT	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	9	47	0.000178	TIGR00482	TIGR00482	cl00015	nicotinate (nicotinamide) nucleotide adenylyltransferase
	non-specific	5	39	0.000485	COG1056	NadR	cl00015	Nicotinamide mononucleotide adenylyltransferase
	non-specific	8	124	0.001859	cd02156	nt_trans	cl00015	nucleotidyl transferase superfamily
	non-specific	8	39	0.007177	cd02166	NMNAT_Archaea	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	5	40	0.007214	PRK00777	PRK00777	cl00015	phosphopantetheine adenylyltransferase
	non-specific	8	39	0.007417	PRK01153	PRK01153	cl00015	nicotinamide-nucleotide adenylyltransferase

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. californicum	non-specific	1	140	6.19e-55	PRK13964	coaD	cl00015	phosphopantetheine adenylyltransferase
	superfamily	1	140	6.19e-55	cl00015	nt_trans superfamily	-	nucleotidyl transferase superfamily
	non-specific	1	134	1.61e-42	COG0669	CoaD	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	1	133	3.31e-37	PRK00168	coaD	cl00015	phosphopantetheine adenylyltransferase
	non-specific	3	133	7.92e-37	cd02163	PPAT	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	3	133	2.43e-33	TIGR01510	coaD_prev_kdtB	cl00015	pantetheine-phosphate adenylyltransferase
	specific	5	135	1.65e-20	pfam01467	CTP_transf_like	cl00015	Cytidylyltransferase-like
	non-specific	3	56	2.09e-10	TIGR00125	cyt_tran_rel	cl00015	cytidyltransferase-like domain
	non-specific	1	66	5.55e-10	PRK00071	nadD	cl00015	nicotinic acid mononucleotide adenylyltransferase
	non-specific	1	66	3.72e-09	COG1057	NadD	cl00015	Nicotinic acid mononucleotide adenylyltransferase
	non-specific	4	66	5.98e-09	cd02165	NMNAT	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	3	133	1.16e-08	cd02039	cytidylyltransferase_like	cl00015	Cytidylyltransferase-like domain
	non-specific	5	66	4.47e-07	TIGR00482	TIGR00482	cl00015	nicotinate (nicotinamide) nucleotide adenylyltransferase
	non-specific	4	58	8.49e-07	PRK05379	PRK05379	cl28366	bifunctional nicotinamide mononucleotide adenylyltransferase
	superfamily	4	58	8.49e-07	cl28366	PRK05379 superfamily	-	bifunctional nicotinamide mononucleotide adenylyltransferase
	non-specific	1	35	1.60e-06	COG1056	NadR	cl00015	Nicotinamide mononucleotide adenylyltransferase
	non-specific	4	65	1.77e-06	cd02168	NMNAT_Nudix	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	5	133	1.90e-06	cd02167	NMNAT_NadR	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	1	71	1.05e-05	PRK07152	nadD	cl28367	putative nicotinate-nucleotide adenylyltransferase
	superfamily	1	71	1.05e-05	cl28367	nadD superfamily	-	putative nicotinate-nucleotide adenylyltransferase
	non-specific	1	33	6.22e-05	PRK00777	PRK00777	cl00015	phosphopantetheine adenylyltransferase
	non-specific	1	133	0.000161	TIGR01526	nadR_NMN_Atrans	cl28365	nicotinamide-nucleotide adenylyltransferase
	superfamily	1	133	0.000161	cl28365	PRK08099 superfamily	-	bifunctional DNA-binding transcriptional repressor/ NMN adenylyltransferase
	non-specific	3	35	0.000733	cd02166	NMNAT_Archaea	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	2	35	0.001521	PRK01153	PRK01153	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	3	46	0.002521	TIGR01527	arch_NMN_Atrans	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	3	56	0.005041	cd02156	nt_trans	cl00015	nucleotidyl transferase superfamily
	non-specific	1	36	0.006027	cd02171	G3P_Cytidylyltransferase	cl00015	glycerol-3-phosphate cytidylyltransferase
M. canis	non-specific	5	139	6.25e-44	PRK13964	coaD	cl00015	phosphopantetheine adenylyltransferase
	superfamily	5	139	6.25e-44	cl00015	nt_trans superfamily	-	nucleotidyl transferase superfamily
	non-specific	5	138	6.62e-40	COG0669	CoaD	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	6	138	1.24e-39	cd02163	PPAT	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	5	138	7.44e-38	PRK00168	coaD	cl00015	phosphopantetheine adenylyltransferase
	non-specific	6	138	2.10e-37	TIGR01510	coaD_prev_kdtB	cl00015	pantetheine-phosphate adenylyltransferase
	specific	8	138	1.27e-19	pfam01467	CTP_transf_like	cl00015	Cytidylyltransferase-like
	non-specific	6	136	4.07e-11	cd02039	cytidylyltransferase_like	cl00015	Cytidylyltransferase-like domain

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. canis	non-specific	6	51	2.76e-10	TIGR00125	cyt_tran_rel	cl00015	cytidyltransferase-like domain
	non-specific	2	47	2.44e-08	COG1056	NadR	cl00015	Nicotinamide mononucleotide adenylyltransferase
	non-specific	6	38	9.48e-08	cd02166	NMNAT_Archaea	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	5	40	2.80e-07	PRK01153	PRK01153	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	7	68	3.94e-07	cd02165	NMNAT	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	1	49	7.06e-06	PRK00071	nadD	cl00015	nicotinic acid mononucleotide adenylyltransferase
	non-specific	6	38	7.70e-06	TIGR01527	arch_NMN_Atrans	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	6	58	3.07e-05	cd02156	nt_trans	cl00015	nucleotidyl transferase superfamily
	non-specific	5	68	4.41e-05	COG1057	NadD	cl00015	Nicotinic acid mononucleotide adenylyltransferase
	non-specific	5	68	0.000197	PRK07152	nadD	cl28367	putative nicotinate-nucleotide adenylyltransferase
	superfamily	5	68	0.000197	cl28367	nadD superfamily	-	putative nicotinate-nucleotide adenylyltransferase
	non-specific	7	47	0.000295	PRK05379	PRK05379	cl28366	bifunctional nicotinamide mononucleotide adenylyltransferase
	superfamily	7	47	0.000295	cl28366	PRK05379 superfamily	-	bifunctional nicotinamide mononucleotide adenylyltransferase
	non-specific	1	57	0.001443	COG1019	CAB4	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	11	36	0.002717	PRK00777	PRK00777	cl00015	phosphopantetheine adenylyltransferase
	non-specific	7	47	0.003878	cd02168	NMNAT_Nudix	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	6	47	0.006711	cd02173	ECT	cl00015	CTP:phosphoethanolamine cytidylyltransferase (ECT)
	non-specific	21	62	0.008447	PTZ00272	PTZ00272	cl27733	heat shock protein 83 kDa (Hsp83)
	superfamily	21	62	0.008447	cl27733	HSP90 superfamily	-	Hsp90 protein
M. capricolum	specific	1	140	2.82e-73	PRK13964	coaD	cl00015	phosphopantetheine adenylyltransferase
	superfamily	1	140	2.82e-73	cl00015	nt_trans superfamily	-	nucleotidyl transferase superfamily
	specific	3	140	1.76e-64	TIGR01510	coaD_prev_kdtB	cl00015	pantetheine-phosphate adenylyltransferase
	non-specific	1	139	2.40e-45	COG0669	CoaD	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	1	135	3.82e-44	PRK00168	coaD	cl00015	phosphopantetheine adenylyltransferase
	non-specific	3	139	1.89e-43	cd02163	PPAT	cl00015	Phosphopantetheine adenylyltransferase
	specific	5	135	5.20e-20	pfam01467	CTP_transf_like	cl00015	Cytidylyltransferase-like
	non-specific	3	63	8.10e-14	TIGR00125	cyt_tran_rel	cl00015	cytidyltransferase-like domain
	non-specific	3	134	2.49e-11	cd02039	cytidylyltransferase_like	cl00015	Cytidylyltransferase-like domain
	non-specific	1	77	2.61e-10	PRK07152	nadD	cl28367	putative nicotinate-nucleotide adenylyltransferase
	superfamily	1	77	2.61e-10	cl28367	nadD superfamily	-	putative nicotinate-nucleotide adenylyltransferase
	non-specific	1	66	1.68e-08	COG1057	NadD	cl00015	Nicotinic acid mononucleotide adenylyltransferase
	non-specific	3	77	1.43e-07	cd02165	NMNAT	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	1	46	2.31e-07	COG1056	NadR	cl00015	Nicotinamide mononucleotide adenylyltransferase
	non-specific	1	136	2.56e-07	PRK00777	PRK00777	cl00015	phosphopantetheine adenylyltransferase
	non-specific	1	66	3.58e-07	PRK00071	nadD	cl00015	nicotinic acid mononucleotide adenylyltransferase
	non-specific	1	136	6.19e-06	COG1019	CAB4	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	1	71	1.09e-05	TIGR01526	nadR_NMN_Atrans	cl28365	nicotinamide-nucleotide adenylyltransferase

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. capricolum	superfamily	1	71	1.09e-05	cl28365	PRK08099 superfamily	-	bifunctional DNA-binding transcriptional repressor/ NMN adenylyltransferase
	non-specific	3	134	1.30e-05	cd02167	NMNAT_NadR	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	4	130	4.30e-05	cd02156	nt_trans	cl00015	nucleotidyl transferase superfamily
	non-specific	4	35	7.26e-05	cd02166	NMNAT_Archaea	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	4	68	0.00016	TIGR01527	arch_NMN_Atrans	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	4	35	0.000417	PRK01153	PRK01153	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	3	35	0.00049	PRK05379	PRK05379	cl28366	bifunctional nicotinamide mononucleotide adenylyltransferase
	superfamily	3	35	0.00049	cl28366	PRK05379 superfamily	-	bifunctional nicotinamide mononucleotide adenylyltransferase
	non-specific	5	60	0.00283	TIGR00482	TIGR00482	cl00015	nicotinate (nicotinamide) nucleotide adenylyltransferase
	non-specific	1	36	0.003701	cd02171	G3P_Cytidylyltransferase	cl00015	glycerol-3-phosphate cytidylyltransferase
	non-specific	8	57	0.004943	cd02170	cytidylyltransferase	cl00015	cytidylyltransferase
	non-specific	1	76	0.004984	PRK08887	PRK08887	cl00015	nicotinic acid mononucleotide adenylyltransferase
M. collis	non-specific	8	145	7.72e-53	PRK13964	coaD	cl00015	phosphopantetheine adenylyltransferase
	superfamily	8	145	7.72e-53	cl00015	nt_trans superfamily	-	nucleotidyl transferase superfamily
	non-specific	6	140	4.38e-43	COG0669	CoaD	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	9	137	1.17e-42	cd02163	PPAT	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	9	139	1.11e-40	TIGR01510	coaD_prev_kdtB	cl00015	pantetheine-phosphate adenylyltransferase
	non-specific	7	137	7.67e-39	PRK00168	coaD	cl00015	phosphopantetheine adenylyltransferase
	non-specific	11	141	2.25e-16	pfam01467	CTP_transf_like	cl00015	Cytidylyltransferase-like
	non-specific	9	69	3.49e-10	TIGR00125	cyt_tran_rel	cl00015	cytidyltransferase-like domain
	non-specific	9	139	2.01e-08	cd02039	cytidylyltransferase_like	cl00015	Cytidylyltransferase-like domain
	non-specific	10	72	1.02e-06	cd02165	NMNAT	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	8	41	1.43e-05	PRK01153	PRK01153	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	9	64	2.26e-05	cd02156	nt_trans	cl00015	nucleotidyl transferase superfamily
	non-specific	9	41	6.88e-05	cd02166	NMNAT_Archaea	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	8	72	8.10e-05	COG1057	NadD	cl00015	Nicotinic acid mononucleotide adenylyltransferase
	non-specific	8	41	9.27e-05	COG1056	NadR	cl00015	Nicotinamide mononucleotide adenylyltransferase
	non-specific	10	57	0.000585	PRK05379	PRK05379	cl28366	bifunctional nicotinamide mononucleotide adenylyltransferase
	superfamily	10	57	0.000585	cl28366	PRK05379 superfamily	-	bifunctional nicotinamide mononucleotide adenylyltransferase
	non-specific	7	48	0.000824	PRK00071	nadD	cl00015	nicotinic acid mononucleotide adenylyltransferase
	non-specific	7	77	0.001049	cd02170	cytidylyltransferase	cl00015	cytidylyltransferase
	non-specific	11	139	0.001357	cd02167	NMNAT_NadR	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	8	72	0.001399	PRK07152	nadD	cl28367	putative nicotinate-nucleotide adenylyltransferase
	superfamily	8	72	0.001399	cl28367	nadD superfamily	-	putative nicotinate-nucleotide adenylyltransferase

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. collis	non-specific	11	45	0.00196	TIGR00482	TIGR00482	cl00015	nicotinate (nicotinamide) nucleotide adenylyltransferase
	non-specific	13	57	0.002052	cd02173	ECT	cl00015	CTP:phosphoethanolamine cytidylyltransferase (ECT)
	non-specific	18	73	0.004238	smart00764	Citrate_ly_lig	cl00015	Citrate lyase ligase C-terminal domain
	non-specific	20	55	0.004797	COG0414	PanC	cl00015	Panthothenate synthetase
	non-specific	18	73	0.005843	cd02169	Citrate_lyase_ligase	cl00015	Citrate lyase ligase
	non-specific	14	49	0.007424	cd02174	CCT	cl00015	CTP:phosphocholine cytidylyltransferase
M. columbinum	non-specific	2	141	2.76e-62	PRK13964	coaD	cl00015	phosphopantetheine adenylyltransferase
	superfamily	2	141	2.76e-62	cl00015	nt_trans superfamily	-	nucleotidyl transferase superfamily
	non-specific	1	141	4.83e-44	COG0669	CoaD	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	4	135	3.24e-38	cd02163	PPAT	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	4	135	1.49e-37	TIGR01510	coaD_prev_kdtB	cl00015	pantetheine-phosphate adenylyltransferase
	non-specific	2	135	4.49e-36	PRK00168	coaD	cl00015	phosphopantetheine adenylyltransferase
	specific	6	137	6.97e-19	pfam01467	CTP_transf_like	cl00015	Cytidylyltransferase-like
	non-specific	4	135	8.19e-11	cd02039	cytidylyltransferase_like	cl00015	Cytidylyltransferase-like domain
	non-specific	4	44	1.63e-10	TIGR00125	cyt_tran_rel	cl00015	cytidyltransferase-like domain
	non-specific	1	59	2.73e-10	COG1057	NadD	cl00015	Nicotinic acid mononucleotide adenylyltransferase
	non-specific	1	68	4.32e-10	PRK00071	nadD	cl00015	nicotinic acid mononucleotide adenylyltransferase
	non-specific	4	42	7.68e-08	cd02165	NMNAT	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	1	36	1.63e-06	COG1056	NadR	cl00015	Nicotinamide mononucleotide adenylyltransferase
	non-specific	2	49	3.02e-06	PRK07152	nadD	cl28367	putative nicotinate-nucleotide adenylyltransferase
	superfamily	2	49	3.02e-06	cl28367	nadD superfamily	-	putative nicotinate-nucleotide adenylyltransferase
	non-specific	4	36	6.72e-06	PRK05379	PRK05379	cl28366	bifunctional nicotinamide mononucleotide adenylyltransferase
	superfamily	4	36	6.72e-06	cl28366	PRK05379 superfamily	-	bifunctional nicotinamide mononucleotide adenylyltransferase
	non-specific	6	61	1.14e-05	TIGR00482	TIGR00482	cl00015	nicotinate (nicotinamide) nucleotide adenylyltransferase
	non-specific	5	94	1.44e-05	cd02166	NMNAT_Archaea	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	4	37	2.74e-05	cd02167	NMNAT_NadR	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	5	65	4.03e-05	cd02156	nt_trans	cl00015	nucleotidyl transferase superfamily
	non-specific	15	40	6.53e-05	pfam02569	Pantoate_ligase	cl00015	Pantoate-beta-alanine ligase
	non-specific	5	64	8.10e-05	TIGR01527	arch_NMN_Atrans	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	1	38	0.000128	PRK08887	PRK08887	cl00015	nicotinic acid mononucleotide adenylyltransferase
	non-specific	2	37	0.000329	TIGR01526	nadR_NMN_Atrans	cl28365	nicotinamide-nucleotide adenylyltransferase
	superfamily	2	37	0.000329	cl28365	PRK08099 superfamily	-	bifunctional DNA-binding transcriptional repressor/ NMN adenylyltransferase
	non-specific	5	36	0.000493	PRK01153	PRK01153	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	1	37	0.00119	cd02170	cytidylyltransferase	cl00015	cytidylyltransferase
	non-specific	15	40	0.001861	cd00560	PanC	cl00015	Pantoate-beta-alanine ligase
	non-specific	1	37	0.00214	cd02171	G3P_Cytidylyltransferase	cl00015	glycerol-3-phosphate cytidylyltransferase
	non-specific	14	40	0.002444	PRK00380	panC	cl00015	pantoate--beta-alanine ligase

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. columbinum	non-specific	1	44	0.00366	COG0615	TagD	cl00015	Glycerol-3-phosphate cytidylyltransferase, cytidylyltransferase family
	non-specific	1	43	0.004032	PRK08099	PRK08099	cl28365	bifunctional DNA-binding transcriptional repressor/ NMN adenylyltransferase
	non-specific	2	42	0.004274	TIGR02199	rfaE_dom_II	cl00015	rfaE bifunctional protein, domain II
	non-specific	2	40	0.005422	PLN02660	PLN02660	cl00015	pantoate--beta-alanine ligase
	non-specific	4	36	0.006818	cd02168	NMNAT_Nudix	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	9	74	0.009055	TIGR01518	g3p_cytidyltrns	cl00015	glycerol-3-phosphate cytidylyltransferase
M. columborale	non-specific	5	141	1.17e-57	PRK13964	coaD	cl00015	phosphopantetheine adenylyltransferase
	superfamily	5	141	1.17e-57	cl00015	nt_trans superfamily	-	nucleotidyl transferase superfamily
	non-specific	5	136	1.15e-41	COG0669	CoaD	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	5	135	6.32e-39	cd02163	PPAT	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	5	135	2.39e-36	PRK00168	coaD	cl00015	phosphopantetheine adenylyltransferase
	non-specific	5	135	4.94e-35	TIGR01510	coaD_prev_kdtB	cl00015	pantetheine-phosphate adenylyltransferase
	specific	7	137	6.56e-23	pfam01467	CTP_transf_like	cl00015	Cytidylyltransferase-like
	non-specific	5	59	3.64e-10	TIGR00125	cyt_tran_rel	cl00015	cytidyltransferase-like domain
	non-specific	5	135	1.53e-08	cd02039	cytidylyltransferase_like	cl00015	Cytidylyltransferase-like domain
	non-specific	1	53	2.08e-07	PRK00071	nadD	cl00015	nicotinic acid mononucleotide adenylyltransferase
	non-specific	5	68	2.50e-07	cd02165	NMNAT	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	5	68	1.42e-06	PRK07152	nadD	cl28367	putative nicotinate-nucleotide adenylyltransferase
	superfamily	5	68	1.42e-06	cl28367	nadD superfamily	-	putative nicotinate-nucleotide adenylyltransferase
	non-specific	1	54	2.17e-05	COG1057	NadD	cl00015	Nicotinic acid mononucleotide adenylyltransferase
	non-specific	6	46	0.000119	cd02156	nt_trans	cl00015	nucleotidyl transferase superfamily
	non-specific	1	38	0.000121	COG1019	CAB4	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	10	38	0.000123	PRK00777	PRK00777	cl00015	phosphopantetheine adenylyltransferase
	non-specific	1	37	0.000197	COG1056	NadR	cl00015	Nicotinamide mononucleotide adenylyltransferase
	non-specific	5	135	0.000343	cd02167	NMNAT_NadR	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	4	135	0.001285	TIGR01526	nadR_NMN_Atrans	cl28365	nicotinamide-nucleotide adenylyltransferase
	superfamily	4	135	0.001285	cl28365	PRK08099 superfamily	-	bifunctional DNA-binding transcriptional repressor/ NMN adenylyltransferase
	non-specific	1	136	0.001697	PRK08099	PRK08099	cl28365	bifunctional DNA-binding transcriptional repressor/ NMN adenylyltransferase
	non-specific	7	62	0.003172	TIGR00482	TIGR00482	cl00015	nicotinate (nicotinamide) nucleotide adenylyltransferase
	non-specific	1	41	0.003186	PRK05379	PRK05379	cl28366	bifunctional nicotinamide mononucleotide adenylyltransferase
	superfamily	1	41	0.003186	cl28366	PRK05379 superfamily	-	bifunctional nicotinamide mononucleotide adenylyltransferase
	non-specific	10	135	0.003564	cd02164	PPAT_CoAS	cl00015	phosphopantetheine adenylyltransferase
	non-specific	5	64	0.006841	PRK01170	PRK01170	cl00866	phosphopantetheine adenylyltransferase
	superfamily	5	64	0.006841	cl00866	NTPase_I-T superfamily	-	Protein of unknown function DUF84

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. columborale	non-specific	6	37	0.008472	cd02166	NMNAT_Archaea	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
M. conjunctivae	specific	10	148	6.49e-74	PRK13964	coaD	cl00015	phosphopantetheine adenylyltransferase
	superfamily	10	148	6.49e-74	cl00015	nt_trans superfamily	-	nucleotidyl transferase superfamily
	non-specific	10	148	6.38e-45	COG0669	CoaD	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	13	140	2.18e-41	cd02163	PPAT	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	13	140	7.70e-39	PRK00168	coaD	cl00015	phosphopantetheine adenylyltransferase
	non-specific	13	142	1.21e-37	TIGR01510	coaD_prev_kdtB	cl00015	pantetheine-phosphate adenylyltransferase
	non-specific	14	144	8.39e-15	pfam01467	CTP_transf_like	cl00015	Cytidylyltransferase-like
	non-specific	13	45	4.87e-10	TIGR00125	cyt_tran_rel	cl00015	cytidyltransferase-like domain
	non-specific	11	75	1.08e-09	PRK00071	nadD	cl00015	nicotinic acid mononucleotide adenylyltransferase
	non-specific	13	75	1.12e-09	cd02165	NMNAT	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	13	142	1.57e-09	cd02039	cytidylyltransferase_like	cl00015	Cytidylyltransferase-like domain
	non-specific	11	75	7.49e-08	COG1057	NadD	cl00015	Nicotinic acid mononucleotide adenylyltransferase
	non-specific	11	81	1.44e-06	PRK07152	nadD	cl28367	putative nicotinate-nucleotide adenylyltransferase
	superfamily	11	81	1.44e-06	cl28367	nadD superfamily	-	putative nicotinate-nucleotide adenylyltransferase
	non-specific	14	64	7.14e-06	TIGR00482	TIGR00482	cl00015	nicotinate (nicotinamide) nucleotide adenylyltransferase
	non-specific	14	75	1.52e-05	cd02167	NMNAT_NadR	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	13	99	1.59e-05	cd02156	nt_trans	cl00015	nucleotidyl transferase superfamily
	non-specific	9	44	3.75e-05	COG1056	NadR	cl00015	Nicotinamide mononucleotide adenylyltransferase
	non-specific	10	45	8.92e-05	cd02170	cytidylyltransferase	cl00015	cytidylyltransferase
	non-specific	22	48	9.34e-05	pfam02569	Pantoate_ligase	cl00015	Pantoate-beta-alanine ligase
	non-specific	13	44	0.000168	PRK05379	PRK05379	cl28366	bifunctional nicotinamide mononucleotide adenylyltransferase
	superfamily	13	44	0.000168	cl28366	PRK05379 superfamily	-	bifunctional nicotinamide mononucleotide adenylyltransferase
	non-specific	11	44	0.001029	PRK01153	PRK01153	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	17	45	0.001618	PRK00777	PRK00777	cl00015	phosphopantetheine adenylyltransferase
	non-specific	13	44	0.001775	cd02166	NMNAT_Archaea	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	22	48	0.002492	COG0414	PanC	cl00015	Panthothenate synthetase
	non-specific	14	80	0.00287	TIGR01526	nadR_NMN_Atrans	cl28365	nicotinamide-nucleotide adenylyltransferase
	superfamily	14	80	0.00287	cl28365	PRK08099 superfamily	-	bifunctional DNA-binding transcriptional repressor/ NMN adenylyltransferase
	non-specific	10	45	0.002877	cd02171	G3P_Cytidylyltransferase	cl00015	glycerol-3-phosphate cytidylyltransferase
	non-specific	22	48	0.003383	PRK00380	panC	cl00015	pantoate--beta-alanine ligase
	non-specific	25	80	0.00395	smart00764	Citrate_ly_lig	cl00015	Citrate lyase ligase C-terminal domain
	non-specific	22	48	0.004388	cd00560	PanC	cl00015	Pantoate-beta-alanine ligase
	non-specific	17	51	0.004435	TIGR02199	rfaE_dom_II	cl00015	rfaE bifunctional protein, domain II
	non-specific	25	80	0.005019	cd02169	Citrate_lyase_ligase	cl00015	Citrate lyase ligase

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. conjunctivae	non-specific	13	44	0.006585	cd02168	NMNAT_Nudix	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	17	96	0.007066	PRK05627	PRK05627	cl27514	bifunctional riboflavin kinase/FMN adenylyltransferase
	superfamily	17	96	0.007066	cl27514	Flavokinase superfamily	-	Riboflavin kinase
	non-specific	2	50	0.009592	PRK08099	PRK08099	cl28365	bifunctional DNA-binding transcriptional repressor/ NMN adenylyltransferase
M. cricetuli	non-specific	3	142	2.92e-62	PRK13964	coaD	cl00015	phosphopantetheine adenylyltransferase
	superfamily	3	142	2.92e-62	cl00015	nt_trans superfamily	-	nucleotidyl transferase superfamily
	non-specific	4	136	8.86e-48	COG0669	CoaD	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	5	135	1.57e-44	cd02163	PPAT	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	5	135	9.74e-44	TIGR01510	coaD_prev_kdtB	cl00015	pantetheine-phosphate adenylyltransferase
	non-specific	4	135	1.86e-43	PRK00168	coaD	cl00015	phosphopantetheine adenylyltransferase
	specific	7	137	1.52e-21	pfam01467	CTP_transf_like	cl00015	Cytidylyltransferase-like
	non-specific	5	65	1.46e-11	TIGR00125	cyt_tran_rel	cl00015	cytidyltransferase-like domain
	non-specific	1	68	1.52e-11	PRK00071	nadD	cl00015	nicotinic acid mononucleotide adenylyltransferase
	non-specific	5	68	1.17e-10	cd02165	NMNAT	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	1	57	7.02e-09	COG1057	NadD	cl00015	Nicotinic acid mononucleotide adenylyltransferase
	non-specific	4	68	2.03e-08	PRK07152	nadD	cl28367	putative nicotinate-nucleotide adenylyltransferase
	superfamily	4	68	2.03e-08	cl28367	nadD superfamily	-	putative nicotinate-nucleotide adenylyltransferase
	non-specific	5	135	4.94e-08	cd02039	cytidylyltransferase_like	cl00015	Cytidylyltransferase-like domain
	non-specific	1	37	3.32e-07	COG1056	NadR	cl00015	Nicotinamide mononucleotide adenylyltransferase
	non-specific	6	48	1.27e-06	cd02156	nt_trans	cl00015	nucleotidyl transferase superfamily
	non-specific	5	135	2.47e-06	cd02167	NMNAT_NadR	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	1	136	1.48e-05	PRK08099	PRK08099	cl28365	bifunctional DNA-binding transcriptional repressor/ NMN adenylyltransferase
	superfamily	1	136	1.48e-05	cl28365	PRK08099 superfamily	-	bifunctional DNA-binding transcriptional repressor/ NMN adenylyltransferase
	non-specific	7	57	1.51e-05	TIGR00482	TIGR00482	cl00015	nicotinate (nicotinamide) nucleotide adenylyltransferase
	non-specific	4	44	2.42e-05	PTZ00308	PTZ00308	cl28626	ethanolamine-phosphate cytidylyltransferase
	superfamily	4	44	2.42e-05	cl28626	PLN02406 superfamily	-	ethanolamine-phosphate cytidylyltransferase
	non-specific	10	38	0.000143	COG0615	TagD	cl00015	Glycerol-3-phosphate cytidylyltransferase, cytidylyltransferase family
	non-specific	10	38	0.000178	cd02170	cytidylyltransferase	cl00015	cytidylyltransferase
	non-specific	6	37	0.000211	PRK01153	PRK01153	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	10	35	0.00061	PRK00777	PRK00777	cl00015	phosphopantetheine adenylyltransferase
	non-specific	7	43	0.000627	cd02173	ECT	cl00015	CTP:phosphoethanolamine cytidylyltransferase (ECT)
	non-specific	6	37	0.000779	cd02166	NMNAT_Archaea	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	4	74	0.001273	TIGR01526	nadR_NMN_Atrans	cl28365	nicotinamide-nucleotide adenylyltransferase
	non-specific	4	38	0.005008	cd02171	G3P_Cytidylyltransferase	cl00015	glycerol-3-phosphate cytidylyltransferase
	non-specific	4	43	0.005313	TIGR02199	rfaE_dom_II	cl00015	rfaE bifunctional protein, domain II

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. cricetuli	non-specific	1	37	0.006315	PRK05379	PRK05379	cl28366	bifunctional nicotinamide mononucleotide adenylyltransferase
	superfamily	1	37	0.006315	cl28366	PRK05379 superfamily	-	bifunctional nicotinamide mononucleotide adenylyltransferase
	non-specific	10	37	0.007769	COG1019	CAB4	cl00015	Phosphopantetheine adenylyltransferase
M. crocodyli	specific	5	146	5.89e-66	PRK13964	coaD	cl00015	phosphopantetheine adenylyltransferase
	superfamily	5	146	5.89e-66	cl00015	nt_trans superfamily	-	nucleotidyl transferase superfamily
	non-specific	6	140	1.71e-39	COG0669	CoaD	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	7	137	2.40e-34	cd02163	PPAT	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	7	146	2.85e-34	TIGR01510	coaD_prev_kdtB	cl00015	pantetheine-phosphate adenylyltransferase
	non-specific	6	137	8.79e-31	PRK00168	coaD	cl00015	phosphopantetheine adenylyltransferase
	non-specific	9	141	8.01e-14	pfam01467	CTP_transf_like	cl00015	Cytidylyltransferase-like
	non-specific	7	58	5.67e-06	TIGR00125	cyt_tran_rel	cl00015	cytidyltransferase-like domain
	non-specific	7	139	0.000547	cd02039	cytidylyltransferase_like	cl00015	Cytidylyltransferase-like domain
	non-specific	8	71	0.008926	cd02165	NMNAT	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
M. felifaucium	non-specific	3	142	8.93e-66	PRK13964	coaD	cl00015	phosphopantetheine adenylyltransferase
	superfamily	3	142	8.93e-66	cl00015	nt_trans superfamily	-	nucleotidyl transferase superfamily
	non-specific	3	136	1.15e-51	COG0669	CoaD	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	5	142	3.50e-47	TIGR01510	coaD_prev_kdtB	cl00015	pantetheine-phosphate adenylyltransferase
	non-specific	5	135	1.94e-46	cd02163	PPAT	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	3	135	3.41e-46	PRK00168	coaD	cl00015	phosphopantetheine adenylyltransferase
	specific	7	135	1.99e-24	pfam01467	CTP_transf_like	cl00015	Cytidylyltransferase-like
	non-specific	5	65	5.00e-14	TIGR00125	cyt_tran_rel	cl00015	cytidyltransferase-like domain
	non-specific	5	135	6.32e-14	cd02039	cytidylyltransferase_like	cl00015	Cytidylyltransferase-like domain
	non-specific	1	37	1.77e-10	COG1056	NadR	cl00015	Nicotinamide mononucleotide adenylyltransferase
	non-specific	1	57	6.34e-09	PRK00071	nadD	cl00015	nicotinic acid mononucleotide adenylyltransferase
	non-specific	6	135	1.07e-08	cd02166	NMNAT_Archaea	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	1	57	3.39e-08	COG1057	NadD	cl00015	Nicotinic acid mononucleotide adenylyltransferase
	non-specific	6	70	1.04e-07	TIGR01527	arch_NMN_Atrans	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	6	37	1.65e-07	PRK01153	PRK01153	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	5	135	3.51e-07	cd02167	NMNAT_NadR	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	5	44	8.59e-07	cd02165	NMNAT	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	3	69	1.08e-06	PRK07152	nadD	cl28367	putative nicotinate-nucleotide adenylyltransferase
	superfamily	3	69	1.08e-06	cl28367	nadD superfamily	-	putative nicotinate-nucleotide adenylyltransferase
	non-specific	7	54	1.57e-06	TIGR00482	TIGR00482	cl00015	nicotinate (nicotinamide) nucleotide adenylyltransferase
	non-specific	3	71	3.90e-06	COG1019	CAB4	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	1	136	3.35e-05	PRK08099	PRK08099	cl28365	bifunctional DNA-binding transcriptional repressor/ NMN adenylyltransferase

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. felifaucium	superfamily	1	136	3.35e-05	cl28365	PRK08099 superfamily	-	bifunctional DNA-binding transcriptional repressor/ NMN adenylyltransferase
	non-specific	5	41	3.99e-05	cd02156	nt_trans	cl00015	nucleotidyl transferase superfamily
	non-specific	3	65	6.86e-05	PRK00777	PRK00777	cl00015	phosphopantetheine adenylyltransferase
	non-specific	1	37	0.00019	PRK05379	PRK05379	cl28366	bifunctional nicotinamide mononucleotide adenylyltransferase
	superfamily	1	37	0.00019	cl28366	PRK05379 superfamily	-	bifunctional nicotinamide mononucleotide adenylyltransferase
	non-specific	3	38	0.000316	TIGR01526	nadR_NMN_Atrans	cl28365	nicotinamide-nucleotide adenylyltransferase
	non-specific	4	44	0.000686	cd02173	ECT	cl00015	CTP:phosphoethanolamine cytidylyltransferase (ECT)
	non-specific	10	57	0.000689	cd02174	CCT	cl00015	CTP:phosphocholine cytidylyltransferase
	non-specific	3	38	0.000828	cd02170	cytidylyltransferase	cl00015	cytidylyltransferase
	non-specific	2	45	0.001162	COG0615	TagD	cl00015	Glycerol-3-phosphate cytidylyltransferase, cytidylyltransferase family
	non-specific	3	43	0.001447	cd02172	RfaE_N	cl00015	N-terminal domain of RfaE
	non-specific	5	27	0.002937	cd02168	NMNAT_Nudix	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	4	56	0.002975	PTZ00308	PTZ00308	cl28626	ethanolamine-phosphate cytidylyltransferase
	superfamily	4	56	0.002975	cl28626	PLN02406 superfamily	-	ethanolamine-phosphate cytidylyltransferase
	non-specific	10	88	0.005345	PRK07143	PRK07143	cl27514	hypothetical protein
	superfamily	10	88	0.005345	cl27514	Flavokinase superfamily	-	Riboflavin kinase
	non-specific	16	48	0.006602	pfam02569	Pantoate_ligase	cl00015	Pantoate-beta-alanine ligase
	non-specific	5	41	0.008476	cd02064	FAD_synthetase_N	cl00015	FAD synthetase, N-terminal domain of the bifunctional enzyme
M. felis	non-specific	2	136	6.63e-51	PRK13964	coaD	cl00015	phosphopantetheine adenylyltransferase
	superfamily	2	136	6.63e-51	cl00015	nt_trans superfamily	-	nucleotidyl transferase superfamily
	non-specific	1	133	4.27e-41	COG0669	CoaD	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	5	133	2.40e-39	TIGR01510	coaD_prev_kdtB	cl00015	pantetheine-phosphate adenylyltransferase
	non-specific	5	131	4.17e-38	cd02163	PPAT	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	5	131	1.97e-34	PRK00168	coaD	cl00015	phosphopantetheine adenylyltransferase
	specific	6	135	1.82e-19	pfam01467	CTP_transf_like	cl00015	Cytidylyltransferase-like
	non-specific	5	48	5.61e-10	TIGR00125	cyt_tran_rel	cl00015	cytidyltransferase-like domain
	non-specific	1	76	3.69e-07	COG1057	NadD	cl00015	Nicotinic acid mononucleotide adenylyltransferase
	non-specific	5	134	3.99e-07	cd02039	cytidylyltransferase_like	cl00015	Cytidylyltransferase-like domain
	non-specific	5	76	5.29e-07	cd02165	NMNAT	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	5	62	1.01e-06	cd02166	NMNAT_Archaea	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	1	76	3.92e-06	PRK00071	nadD	cl00015	nicotinic acid mononucleotide adenylyltransferase
	non-specific	1	46	1.13e-05	COG1056	NadR	cl00015	Nicotinamide mononucleotide adenylyltransferase
	non-specific	5	36	1.17e-05	PRK05379	PRK05379	cl28366	bifunctional nicotinamide mononucleotide adenylyltransferase
	superfamily	5	36	1.17e-05	cl28366	PRK05379 superfamily	-	bifunctional nicotinamide mononucleotide adenylyltransferase

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. felis	non-specific	6	76	2.15e-05	TIGR00482	TIGR00482	cl00015	nicotinate (nicotinamide) nucleotide adenylyltransferase
	non-specific	5	46	2.27e-05	PRK01153	PRK01153	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	9	48	1.00e-04	cd02174	CCT	cl00015	CTP:phosphocholine cytidylyltransferase
	non-specific	5	68	0.000101	TIGR01527	arch_NMN_Atrans	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	5	66	0.000241	PRK07152	nadD	cl28367	putative nicotinate-nucleotide adenylyltransferase
	superfamily	5	66	0.000241	cl28367	nadD superfamily	-	putative nicotinate-nucleotide adenylyltransferase
	non-specific	5	36	0.000349	cd02168	NMNAT_Nudix	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	1	37	0.001005	PRK08887	PRK08887	cl00015	nicotinic acid mononucleotide adenylyltransferase
	non-specific	5	40	0.002275	cd02156	nt_trans	cl00015	nucleotidyl transferase superfamily
	non-specific	9	86	0.006609	COG0615	TagD	cl00015	Glycerol-3-phosphate cytidylyltransferase, cytidylyltransferase family
M. fermentans	non-specific	3	141	7.03e-63	PRK13964	coaD	cl00015	phosphopantetheine adenylyltransferase
	superfamily	3	141	7.03e-63	cl00015	nt_trans superfamily	-	nucleotidyl transferase superfamily
	non-specific	1	140	8.44e-50	COG0669	CoaD	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	4	134	3.86e-47	cd02163	PPAT	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	3	134	3.77e-46	PRK00168	coaD	cl00015	phosphopantetheine adenylyltransferase
	non-specific	4	134	2.60e-42	TIGR01510	coaD_prev_kdtB	cl00015	pantetheine-phosphate adenylyltransferase
	specific	6	136	5.06e-23	pfam01467	CTP_transf_like	cl00015	Cytidylyltransferase-like
	non-specific	4	56	2.61e-14	TIGR00125	cyt_tran_rel	cl00015	cytidyltransferase-like domain
	non-specific	1	68	4.23e-12	COG1057	NadD	cl00015	Nicotinic acid mononucleotide adenylyltransferase
	non-specific	4	134	6.66e-12	cd02039	cytidylyltransferase_like	cl00015	Cytidylyltransferase-like domain
	non-specific	1	67	2.65e-11	PRK00071	nadD	cl00015	nicotinic acid mononucleotide adenylyltransferase
	non-specific	5	67	8.01e-11	cd02165	NMNAT	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	3	36	5.60e-09	COG1056	NadR	cl00015	Nicotinamide mononucleotide adenylyltransferase
	non-specific	9	134	1.07e-08	cd02167	NMNAT_NadR	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	6	61	2.32e-08	TIGR00482	TIGR00482	cl00015	nicotinate (nicotinamide) nucleotide adenylyltransferase
	non-specific	3	36	4.28e-08	PRK01153	PRK01153	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	4	63	6.34e-08	cd02166	NMNAT_Archaea	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	4	56	9.89e-08	cd02156	nt_trans	cl00015	nucleotidyl transferase superfamily
	non-specific	3	61	2.69e-07	PRK07152	nadD	cl28367	putative nicotinate-nucleotide adenylyltransferase
	superfamily	3	61	2.69e-07	cl28367	nadD superfamily	-	putative nicotinate-nucleotide adenylyltransferase
	non-specific	4	67	2.77e-07	TIGR01527	arch_NMN_Atrans	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	5	61	8.30e-06	PRK05379	PRK05379	cl28366	bifunctional nicotinamide mononucleotide adenylyltransferase
	superfamily	5	61	8.30e-06	cl28366	PRK05379 superfamily	-	bifunctional nicotinamide mononucleotide adenylyltransferase
	non-specific	9	73	2.49e-05	TIGR01526	nadR_NMN_Atrans	cl28365	nicotinamide-nucleotide adenylyltransferase
	superfamily	9	73	2.49e-05	cl28365	PRK08099 superfamily	-	bifunctional DNA-binding transcriptional repressor/ NMN adenylyltransferase

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. fermentans	non-specific	9	56	3.99e-05	COG0615	TagD	cl00015	Glycerol-3-phosphate cytidylyltransferase, cytidylyltransferase family
	non-specific	9	37	4.05e-05	cd02170	cytidylyltransferase	cl00015	cytidylyltransferase
	non-specific	5	78	9.88e-05	cd02168	NMNAT_Nudix	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	3	42	0.000254	TIGR02199	rfaE_dom_II	cl00015	rfaE bifunctional protein, domain II
	non-specific	9	56	0.000745	cd02174	CCT	cl00015	CTP:phosphocholine cytidylyltransferase
	non-specific	15	40	0.000758	pfam02569	Pantoate_ligase	cl00015	Pantoate-beta-alanine ligase
	non-specific	9	57	0.000849	TIGR01518	g3p_cytidyltrns	cl00015	glycerol-3-phosphate cytidylyltransferase
	non-specific	8	55	0.000992	PTZ00308	PTZ00308	cl28626	ethanolamine-phosphate cytidylyltransferase
	superfamily	8	55	0.000992	cl28626	PLN02406 superfamily	-	ethanolamine-phosphate cytidylyltransferase
	non-specific	3	64	0.001208	cd02171	G3P_Cytidylyltransferase	cl00015	glycerol-3-phosphate cytidylyltransferase
	non-specific	9	42	0.002815	COG2870	RfaE	cl28454	ADP-heptose synthase, bifunctional sugar kinase/adenylyltransferase
	superfamily	9	42	0.002815	cl28454	RfaE superfamily	-	ADP-heptose synthase, bifunctional sugar kinase/adenylyltransferase
	non-specific	9	135	0.00309	PRK08099	PRK08099	cl28365	bifunctional DNA-binding transcriptional repressor/ NMN adenylyltransferase
	non-specific	3	69	0.00316	PRK00777	PRK00777	cl00015	phosphopantetheine adenylyltransferase
	non-specific	9	36	0.005021	PRK11316	PRK11316	cl28454	bifunctional heptose 7-phosphate kinase/heptose 1-phosphate adenyltransferase
	non-specific	8	34	0.007501	cd02173	ECT	cl00015	CTP:phosphoethanolamine cytidylyltransferase (ECT)
M. gallinaceum	non-specific	5	143	9.54e-64	PRK13964	coaD	cl00015	phosphopantetheine adenylyltransferase
	superfamily	5	143	9.54e-64	cl00015	nt_trans superfamily	-	nucleotidyl transferase superfamily
	non-specific	3	145	1.96e-45	COG0669	CoaD	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	6	145	1.02e-41	TIGR01510	coaD_prev_kdtB	cl00015	pantetheine-phosphate adenylyltransferase
	non-specific	5	145	9.37e-39	PRK00168	coaD	cl00015	phosphopantetheine adenylyltransferase
	non-specific	6	136	1.86e-38	cd02163	PPAT	cl00015	Phosphopantetheine adenylyltransferase
	specific	8	138	1.46e-19	pfam01467	CTP_transf_like	cl00015	Cytidylyltransferase-like
	non-specific	6	136	2.50e-11	cd02039	cytidylyltransferase_like	cl00015	Cytidylyltransferase-like domain
	non-specific	6	38	2.53e-09	TIGR00125	cyt_tran_rel	cl00015	cytidyltransferase-like domain
	non-specific	7	69	3.46e-08	cd02165	NMNAT	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	5	69	9.23e-08	PRK00071	nadD	cl00015	nicotinic acid mononucleotide adenylyltransferase
	non-specific	6	95	4.66e-06	cd02156	nt_trans	cl00015	nucleotidyl transferase superfamily
	non-specific	8	69	7.41e-06	TIGR00482	TIGR00482	cl00015	nicotinate (nicotinamide) nucleotide adenylyltransferase
	non-specific	5	64	1.59e-05	COG1057	NadD	cl00015	Nicotinic acid mononucleotide adenylyltransferase
	non-specific	8	48	7.01e-05	cd02167	NMNAT_NadR	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	8	44	0.00012	cd02174	CCT	cl00015	CTP:phosphocholine cytidylyltransferase
	non-specific	5	39	0.000131	cd02170	cytidylyltransferase	cl00015	cytidylyltransferase
	non-specific	5	38	0.000193	COG1056	NadR	cl00015	Nicotinamide mononucleotide adenylyltransferase
	non-specific	6	44	0.0002	cd02173	ECT	cl00015	CTP:phosphoethanolamine cytidylyltransferase (ECT)

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. gallinaceum	non-specific	3	44	0.000461	PTZ00308	PTZ00308	cl28626	ethanolamine-phosphate cytidylyltransferase
	superfamily	3	44	0.000461	cl28626	PLN02406 superfamily	-	ethanolamine-phosphate cytidylyltransferase
	non-specific	5	69	0.000576	PRK07152	nadD	cl28367	putative nicotinate-nucleotide adenylyltransferase
	superfamily	5	69	0.000576	cl28367	nadD superfamily	-	putative nicotinate-nucleotide adenylyltransferase
	non-specific	5	139	0.000995	PRK00777	PRK00777	cl00015	phosphopantetheine adenylyltransferase
	non-specific	5	39	0.001682	cd02171	G3P_Cytidylyltransferase	cl00015	glycerol-3-phosphate cytidylyltransferase
	non-specific	5	39	0.002016	COG0615	TagD	cl00015	Glycerol-3-phosphate cytidylyltransferase, cytidylyltransferase family
	non-specific	8	48	0.002371	TIGR01526	nadR_NMN_Atrans	cl28365	nicotinamide-nucleotide adenylyltransferase
	superfamily	8	48	0.002371	cl28365	PRK08099 superfamily	-	bifunctional DNA-binding transcriptional repressor/ NMN adenylyltransferase
	non-specific	1	66	0.006453	COG1019	CAB4	cl00015	Phosphopantetheine adenylyltransferase
M. gallinarum	non-specific	2	141	9.93e-64	PRK13964	coaD	cl00015	phosphopantetheine adenylyltransferase
	superfamily	2	141	9.93e-64	cl00015	nt_trans superfamily	-	nucleotidyl transferase superfamily
	non-specific	1	140	1.36e-47	COG0669	CoaD	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	4	141	3.25e-44	TIGR01510	coaD_prev_kdtB	cl00015	pantetheine-phosphate adenylyltransferase
	non-specific	3	132	8.06e-42	PRK00168	coaD	cl00015	phosphopantetheine adenylyltransferase
	non-specific	4	132	9.66e-42	cd02163	PPAT	cl00015	Phosphopantetheine adenylyltransferase
	specific	6	132	1.87e-20	pfam01467	CTP_transf_like	cl00015	Cytidylyltransferase-like
	non-specific	4	64	8.98e-14	TIGR00125	cyt_tran_rel	cl00015	cytidyltransferase-like domain
	non-specific	3	67	3.27e-11	PRK07152	nadD	cl28367	putative nicotinate-nucleotide adenylyltransferase
	superfamily	3	67	3.27e-11	cl28367	nadD superfamily	-	putative nicotinate-nucleotide adenylyltransferase
	non-specific	1	56	2.69e-10	COG1057	NadD	cl00015	Nicotinic acid mononucleotide adenylyltransferase
	non-specific	1	36	2.25e-09	COG1056	NadR	cl00015	Nicotinamide mononucleotide adenylyltransferase
	non-specific	4	134	3.06e-09	cd02039	cytidylyltransferase_like	cl00015	Cytidylyltransferase-like domain
	non-specific	4	67	3.78e-09	cd02165	NMNAT	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	1	53	1.54e-08	PRK00071	nadD	cl00015	nicotinic acid mononucleotide adenylyltransferase
	non-specific	5	56	6.30e-08	PRK05379	PRK05379	cl28366	bifunctional nicotinamide mononucleotide adenylyltransferase
	superfamily	5	56	6.30e-08	cl28366	PRK05379 superfamily	-	bifunctional nicotinamide mononucleotide adenylyltransferase
	non-specific	6	56	6.69e-08	TIGR00482	TIGR00482	cl00015	nicotinate (nicotinamide) nucleotide adenylyltransferase
	non-specific	5	36	8.94e-07	PRK01153	PRK01153	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	4	91	9.39e-07	cd02156	nt_trans	cl00015	nucleotidyl transferase superfamily
	non-specific	5	36	1.10e-06	cd02166	NMNAT_Archaea	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	1	57	3.21e-06	cd02170	cytidylyltransferase	cl00015	cytidylyltransferase
	non-specific	2	84	4.94e-06	TIGR02199	rfaE_dom_II	cl00015	rfaE bifunctional protein, domain II
	non-specific	4	60	5.26e-06	cd02168	NMNAT_Nudix	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	1	72	1.88e-05	COG0615	TagD	cl00015	Glycerol-3-phosphate cytidylyltransferase, cytidylyltransferase family

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. gallinarum	non-specific	4	43	2.66e-05	cd02167	NMNAT_NadR	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	5	67	3.72e-05	TIGR01527	arch_NMN_Atrans	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	1	43	5.45e-05	PRK08099	PRK08099	cl28365	bifunctional DNA-binding transcriptional repressor/ NMN adenylyltransferase
	superfamily	1	43	5.45e-05	cl28365	PRK08099 superfamily	-	bifunctional DNA-binding transcriptional repressor/ NMN adenylyltransferase
	non-specific	2	37	7.50e-05	cd02171	G3P_Cytidylyltransferase	cl00015	glycerol-3-phosphate cytidylyltransferase
	non-specific	3	37	8.20e-05	TIGR01526	nadR_NMN_Atrans	cl28365	nicotinamide-nucleotide adenylyltransferase
	non-specific	2	36	0.000151	PTZ00308	PTZ00308	cl28626	ethanolamine-phosphate cytidylyltransferase
	superfamily	2	36	0.000151	cl28626	PLN02406 superfamily	-	ethanolamine-phosphate cytidylyltransferase
	non-specific	3	64	0.00023	PRK01170	PRK01170	cl00866	phosphopantetheine adenylyltransferase
	superfamily	3	64	0.00023	cl00866	NTPase_I-T superfamily	-	Protein of unknown function DUF84
	non-specific	14	40	0.000247	pfam02569	Pantoate_ligase	cl00015	Pantoate-beta-alanine ligase
	non-specific	1	34	0.000283	PRK00777	PRK00777	cl00015	phosphopantetheine adenylyltransferase
	non-specific	2	36	0.000525	PRK11316	PRK11316	cl28454	bifunctional heptose 7-phosphate kinase/heptose 1-phosphate adenyltransferase
	superfamily	2	36	0.000525	cl28454	RfaE superfamily	-	ADP-heptose synthase, bifunctional sugar kinase/adenylyltransferase
	non-specific	4	40	0.000654	pfam06574	FAD_syn	cl00015	FAD synthetase
	non-specific	9	64	0.000754	COG1019	CAB4	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	4	40	0.000868	PRK05627	PRK05627	cl27514	bifunctional riboflavin kinase/FMN adenylyltransferase
	superfamily	4	40	0.000868	cl27514	Flavokinase superfamily	-	Riboflavin kinase
	non-specific	6	34	0.000934	cd02173	ECT	cl00015	CTP:phosphoethanolamine cytidylyltransferase (ECT)
	non-specific	4	40	0.001299	cd02064	FAD_synthetase_N	cl00015	FAD synthetase, N-terminal domain of the bifunctional enzyme
	non-specific	9	42	0.001831	COG2870	RfaE	cl28454	ADP-heptose synthase, bifunctional sugar kinase/adenylyltransferase
	non-specific	2	79	0.003548	PRK07143	PRK07143	cl27514	hypothetical protein
	non-specific	9	56	0.003727	cd02174	CCT	cl00015	CTP:phosphocholine cytidylyltransferase
	non-specific	9	64	0.004673	cd02164	PPAT_CoAS	cl00015	phosphopantetheine adenylyltransferase
	non-specific	9	54	0.006477	TIGR01518	g3p_cytidyltrns	cl00015	glycerol-3-phosphate cytidylyltransferase
	non-specific	2	42	0.009744	cd02172	RfaE_N	cl00015	N-terminal domain of RfaE
M. iners	non-specific	2	141	1.65e-55	PRK13964	coaD	cl00015	phosphopantetheine adenylyltransferase
	superfamily	2	141	1.65e-55	cl00015	nt_trans superfamily	-	nucleotidyl transferase superfamily
	non-specific	1	139	3.63e-41	COG0669	CoaD	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	2	132	9.44e-39	PRK00168	coaD	cl00015	phosphopantetheine adenylyltransferase
	non-specific	4	132	2.28e-38	cd02163	PPAT	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	4	136	7.71e-33	TIGR01510	coaD_prev_kdtB	cl00015	pantetheine-phosphate adenylyltransferase
	specific	6	141	3.41e-17	pfam01467	CTP_transf_like	cl00015	Cytidylyltransferase-like
	non-specific	1	67	7.10e-12	COG1057	NadD	cl00015	Nicotinic acid mononucleotide adenylyltransferase
	non-specific	4	64	1.05e-11	TIGR00125	cyt_tran_rel	cl00015	cytidyltransferase-like domain

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. iners	non-specific	1	135	1.44e-11	PRK08099	PRK08099	cl28365	bifunctional DNA-binding transcriptional repressor/ NMN adenylyltransferase
	superfamily	1	135	1.44e-11	cl28365	PRK08099 superfamily	-	bifunctional DNA-binding transcriptional repressor/ NMN adenylyltransferase
	non-specific	1	67	4.81e-11	PRK00071	nadD	cl00015	nicotinic acid mononucleotide adenylyltransferase
	non-specific	4	134	2.44e-10	cd02167	NMNAT_NadR	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	2	67	4.93e-10	PRK07152	nadD	cl28367	putative nicotinate-nucleotide adenylyltransferase
	superfamily	2	67	4.93e-10	cl28367	nadD superfamily	-	putative nicotinate-nucleotide adenylyltransferase
	non-specific	4	134	2.26e-09	cd02039	cytidylyltransferase_like	cl00015	Cytidylyltransferase-like domain
	non-specific	6	67	2.80e-09	TIGR00482	TIGR00482	cl00015	nicotinate (nicotinamide) nucleotide adenylyltransferase
	non-specific	1	63	4.74e-09	COG1056	NadR	cl00015	Nicotinamide mononucleotide adenylyltransferase
	non-specific	4	67	8.32e-09	cd02165	NMNAT	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	2	36	3.02e-06	TIGR01526	nadR_NMN_Atrans	cl28365	nicotinamide-nucleotide adenylyltransferase
	non-specific	5	63	3.59e-06	PRK01153	PRK01153	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	6	91	3.62e-06	cd02156	nt_trans	cl00015	nucleotidyl transferase superfamily
	non-specific	5	63	1.28e-05	cd02166	NMNAT_Archaea	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	4	36	0.000109	PRK05379	PRK05379	cl28366	bifunctional nicotinamide mononucleotide adenylyltransferase
	superfamily	4	36	0.000109	cl28366	PRK05379 superfamily	-	bifunctional nicotinamide mononucleotide adenylyltransferase
	non-specific	9	37	0.000179	PRK00777	PRK00777	cl00015	phosphopantetheine adenylyltransferase
	non-specific	1	37	0.000229	cd02170	cytidylyltransferase	cl00015	cytidylyltransferase
	non-specific	5	61	0.000865	TIGR01527	arch_NMN_Atrans	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	3	30	0.001251	PRK06973	PRK06973	cl00015	nicotinic acid mononucleotide adenylyltransferase
	non-specific	9	40	0.002132	pfam06574	FAD_syn	cl00015	FAD synthetase
	non-specific	9	37	0.003561	COG1019	CAB4	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	9	40	0.003964	PRK05627	PRK05627	cl27514	bifunctional riboflavin kinase/FMN adenylyltransferase
	superfamily	9	40	0.003964	cl27514	Flavokinase superfamily	-	Riboflavin kinase
	non-specific	3	42	0.004603	PTZ00308	PTZ00308	cl28626	ethanolamine-phosphate cytidylyltransferase
	superfamily	3	42	0.004603	cl28626	PLN02406 superfamily	-	ethanolamine-phosphate cytidylyltransferase
	non-specific	1	37	0.004662	cd02171	G3P_Cytidylyltransferase	cl00015	glycerol-3-phosphate cytidylyltransferase
	non-specific	1	36	0.004933	COG0615	TagD	cl00015	Glycerol-3-phosphate cytidylyltransferase, cytidylyltransferase family
	non-specific	2	40	0.006074	PLN02660	PLN02660	cl00015	pantoate--beta-alanine ligase
M. iowae	non-specific	8	141	5.58e-31	PRK13964	coaD	cl00015	phosphopantetheine adenylyltransferase
	superfamily	8	141	5.58e-31	cl00015	nt_trans superfamily	-	nucleotidyl transferase superfamily
	non-specific	8	136	1.04e-28	COG0669	CoaD	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	8	140	9.33e-27	TIGR01510	coaD_prev_kdtB	cl00015	pantetheine-phosphate adenylyltransferase
	non-specific	8	136	7.25e-26	cd02163	PPAT	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	8	136	1.01e-23	PRK00168	coaD	cl00015	phosphopantetheine adenylyltransferase

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. iowae	non-specific	9	138	2.07e-07	pfam01467	CTP_transf_like	cl00015	Cytidylyltransferase-like
	non-specific	8	93	0.001153	cd02156	nt_trans	cl00015	nucleotidyl transferase superfamily
	non-specific	8	64	0.001154	TIGR00125	cyt_tran_rel	cl00015	cytidyltransferase-like domain
	non-specific	8	136	0.001185	cd02039	cytidylyltransferase_like	cl00015	Cytidylyltransferase-like domain
M. leachii	specific	1	140	3.91e-73	PRK13964	coaD	cl00015	phosphopantetheine adenylyltransferase
	superfamily	1	140	3.91e-73	cl00015	nt_trans superfamily	-	nucleotidyl transferase superfamily
	non-specific	3	140	4.52e-64	TIGR01510	coaD_prev_kdtB	cl00015	pantetheine-phosphate adenylyltransferase
	non-specific	1	139	8.52e-46	COG0669	CoaD	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	1	135	4.75e-44	PRK00168	coaD	cl00015	phosphopantetheine adenylyltransferase
	non-specific	3	139	1.50e-43	cd02163	PPAT	cl00015	Phosphopantetheine adenylyltransferase
	specific	5	135	8.49e-21	pfam01467	CTP_transf_like	cl00015	Cytidylyltransferase-like
	non-specific	4	63	1.33e-13	TIGR00125	cyt_tran_rel	cl00015	cytidyltransferase-like domain
	non-specific	4	134	6.52e-12	cd02039	cytidylyltransferase_like	cl00015	Cytidylyltransferase-like domain
	non-specific	1	77	1.09e-09	PRK07152	nadD	cl28367	putative nicotinate-nucleotide adenylyltransferase
	superfamily	1	77	1.09e-09	cl28367	nadD superfamily	-	putative nicotinate-nucleotide adenylyltransferase
	non-specific	1	66	1.65e-07	COG1057	NadD	cl00015	Nicotinic acid mononucleotide adenylyltransferase
	non-specific	3	77	4.15e-07	cd02165	NMNAT	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	1	136	5.19e-07	PRK00777	PRK00777	cl00015	phosphopantetheine adenylyltransferase
	non-specific	1	46	1.35e-06	COG1056	NadR	cl00015	Nicotinamide mononucleotide adenylyltransferase
	non-specific	1	66	1.57e-06	PRK00071	nadD	cl00015	nicotinic acid mononucleotide adenylyltransferase
	non-specific	1	136	6.25e-06	COG1019	CAB4	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	3	130	8.03e-06	cd02156	nt_trans	cl00015	nucleotidyl transferase superfamily
	non-specific	3	134	2.76e-05	cd02167	NMNAT_NadR	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	4	68	4.23e-05	cd02166	NMNAT_Archaea	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	1	71	7.22e-05	TIGR01526	nadR_NMN_Atrans	cl28365	nicotinamide-nucleotide adenylyltransferase
	superfamily	1	71	7.22e-05	cl28365	PRK08099 superfamily	-	bifunctional DNA-binding transcriptional repressor/ NMN adenylyltransferase
	non-specific	4	68	0.00016	TIGR01527	arch_NMN_Atrans	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	4	35	0.000363	PRK01153	PRK01153	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	4	35	0.00079	PRK05379	PRK05379	cl28366	bifunctional nicotinamide mononucleotide adenylyltransferase
	superfamily	4	35	0.00079	cl28366	PRK05379 superfamily	-	bifunctional nicotinamide mononucleotide adenylyltransferase
	non-specific	5	66	0.001253	TIGR00482	TIGR00482	cl00015	nicotinate (nicotinamide) nucleotide adenylyltransferase
	non-specific	1	89	0.00256	cd02170	cytidylyltransferase	cl00015	cytidylyltransferase
	non-specific	1	36	0.002969	cd02171	G3P_Cytidylyltransferase	cl00015	glycerol-3-phosphate cytidylyltransferase
M. leonicaptivi	non-specific	2	134	1.38e-46	COG0669	CoaD	cl00015	Phosphopantetheine adenylyltransferase
	superfamily	2	134	1.38e-46	cl00015	nt_trans superfamily	-	nucleotidyl transferase superfamily
	non-specific	5	134	7.67e-46	cd02163	PPAT	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	3	134	3.09e-44	PRK00168	coaD	cl00015	phosphopantetheine adenylyltransferase

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. leonicaptivi	non-specific	3	137	4.65e-42	PRK13964	coaD	cl00015	phosphopantetheine adenylyltransferase
	non-specific	5	134	2.30e-41	TIGR01510	coaD_prev_kdtB	cl00015	pantetheine-phosphate adenylyltransferase
	specific	8	136	9.01e-24	pfam01467	CTP_transf_like	cl00015	Cytidylyltransferase-like
	non-specific	5	135	1.45e-12	cd02039	cytidylyltransferase_like	cl00015	Cytidylyltransferase-like domain
	non-specific	5	48	1.59e-11	TIGR00125	cyt_tran_rel	cl00015	cytidyltransferase-like domain
	non-specific	1	46	2.03e-10	PRK00071	nadD	cl00015	nicotinic acid mononucleotide adenylyltransferase
	non-specific	1	48	5.65e-09	COG1057	NadD	cl00015	Nicotinic acid mononucleotide adenylyltransferase
	non-specific	6	48	9.93e-08	PRK01153	PRK01153	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	1	48	2.84e-07	COG1056	NadR	cl00015	Nicotinamide mononucleotide adenylyltransferase
	non-specific	6	48	5.20e-07	cd02166	NMNAT_Archaea	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	1	86	8.65e-07	COG0615	TagD	cl00015	Glycerol-3-phosphate cytidylyltransferase, cytidylyltransferase family
	non-specific	5	46	1.25e-06	cd02165	NMNAT	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	10	137	1.71e-06	COG1019	CAB4	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	4	48	5.31e-06	PRK05379	PRK05379	cl28366	bifunctional nicotinamide mononucleotide adenylyltransferase
	superfamily	4	48	5.31e-06	cl28366	PRK05379 superfamily	-	bifunctional nicotinamide mononucleotide adenylyltransferase
	non-specific	3	88	1.99e-05	PRK07152	nadD	cl28367	putative nicotinate-nucleotide adenylyltransferase
	superfamily	3	88	1.99e-05	cl28367	nadD superfamily	-	putative nicotinate-nucleotide adenylyltransferase
	non-specific	5	48	2.27e-05	cd02168	NMNAT_Nudix	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	1	38	3.74e-05	cd02170	cytidylyltransferase	cl00015	cytidylyltransferase
	non-specific	10	137	5.81e-05	PRK00777	PRK00777	cl00015	phosphopantetheine adenylyltransferase
	non-specific	8	49	6.14e-05	TIGR00482	TIGR00482	cl00015	nicotinate (nicotinamide) nucleotide adenylyltransferase
	non-specific	6	132	9.09e-05	cd02156	nt_trans	cl00015	nucleotidyl transferase superfamily
	non-specific	8	69	0.00066	TIGR01527	arch_NMN_Atrans	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	1	135	0.00071	PRK08099	PRK08099	cl28365	bifunctional DNA-binding transcriptional repressor/ NMN adenylyltransferase
	superfamily	1	135	0.00071	cl28365	PRK08099 superfamily	-	bifunctional DNA-binding transcriptional repressor/ NMN adenylyltransferase
	non-specific	10	137	0.001165	PRK01170	PRK01170	cl00866	phosphopantetheine adenylyltransferase
	superfamily	10	137	0.001165	cl00866	NTPase_I-T superfamily	-	Protein of unknown function DUF84
	non-specific	3	57	0.001427	TIGR01526	nadR_NMN_Atrans	cl28365	nicotinamide-nucleotide adenylyltransferase
	non-specific	12	38	0.005519	pfam05636	HIGH_NTase1	cl27012	HIGH Nucleotidyl Transferase
	superfamily	12	38	0.005519	cl27012	HIGH_NTase1 superfamily	-	HIGH Nucleotidyl Transferase
	non-specific	2	37	0.006823	cd02173	ECT	cl00015	CTP:phosphoethanolamine cytidylyltransferase (ECT)
	non-specific	7	135	0.006888	cd02164	PPAT_CoAS	cl00015	phosphopantetheine adenylyltransferase
	non-specific	2	80	0.007169	cd02169	Citrate_lyase_ligase	cl00015	Citrate lyase ligase
	non-specific	10	38	0.008445	cd09286	NMNAT_Eukarya	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
M. lipofaciens	non-specific	1	140	9.74e-63	PRK13964	coaD	cl00015	phosphopantetheine adenylyltransferase

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. lipofaciens	superfamily	1	140	9.74e-63	cl00015	nt_trans superfamily	-	nucleotidyl transferase superfamily
	non-specific	1	134	4.02e-43	COG0669	CoaD	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	3	133	1.48e-38	cd02163	PPAT	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	3	133	1.16e-37	TIGR01510	coaD_prev_kdtB	cl00015	pantetheine-phosphate adenylyltransferase
	non-specific	1	133	2.28e-37	PRK00168	coaD	cl00015	phosphopantetheine adenylyltransferase
	non-specific	5	135	1.17e-15	pfam01467	CTP_transf_like	cl00015	Cytidylyltransferase-like
	non-specific	3	63	6.62e-12	TIGR00125	cyt_tran_rel	cl00015	cytidyltransferase-like domain
	non-specific	1	66	1.19e-09	COG1057	NadD	cl00015	Nicotinic acid mononucleotide adenylyltransferase
	non-specific	3	133	4.38e-08	cd02039	cytidylyltransferase_like	cl00015	Cytidylyltransferase-like domain
	non-specific	3	90	7.99e-08	cd02156	nt_trans	cl00015	nucleotidyl transferase superfamily
	non-specific	5	133	1.98e-07	cd02167	NMNAT_NadR	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	4	66	2.34e-07	cd02165	NMNAT	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	1	66	1.08e-06	PRK07152	nadD	cl28367	putative nicotinate-nucleotide adenylyltransferase
	superfamily	1	66	1.08e-06	cl28367	nadD superfamily	-	putative nicotinate-nucleotide adenylyltransferase
	non-specific	1	35	1.39e-06	COG1056	NadR	cl00015	Nicotinamide mononucleotide adenylyltransferase
	non-specific	1	133	1.77e-05	cd02170	cytidylyltransferase	cl00015	cytidylyltransferase
	non-specific	1	66	2.74e-05	PRK00071	nadD	cl00015	nicotinic acid mononucleotide adenylyltransferase
	non-specific	5	52	3.21e-05	TIGR00482	TIGR00482	cl00015	nicotinate (nicotinamide) nucleotide adenylyltransferase
	non-specific	1	43	8.20e-05	COG0615	TagD	cl00015	Glycerol-3-phosphate cytidylyltransferase, cytidylyltransferase family
	non-specific	13	39	9.36e-05	pfam02569	Pantoate_ligase	cl00015	Pantoate-beta-alanine ligase
	non-specific	1	133	0.000121	TIGR01526	nadR_NMN_Atrans	cl28365	nicotinamide-nucleotide adenylyltransferase
	superfamily	1	133	0.000121	cl28365	PRK08099 superfamily	-	bifunctional DNA-binding transcriptional repressor/ NMN adenylyltransferase
	non-specific	2	35	0.000121	PRK01153	PRK01153	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	1	72	0.000135	cd02171	G3P_Cytidylyltransferase	cl00015	glycerol-3-phosphate cytidylyltransferase
	non-specific	34	94	0.000149	PRK10422	PRK10422	cl10013	lipopolysaccharide core biosynthesis protein
	superfamily	34	94	0.000149	cl10013	Glycosyltransferase_GTB_type superfamily	-	Glycosyltransferases catalyze the transfer of sugar moieties from activated donor molecules to specific acceptor molecules, forming glycosidic bonds.
	non-specific	5	61	0.000181	PLN02388	PLN02388	cl00015	phosphopantetheine adenylyltransferase
	non-specific	3	35	0.000184	cd02166	NMNAT_Archaea	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	1	63	0.000198	PRK00777	PRK00777	cl00015	phosphopantetheine adenylyltransferase
	non-specific	4	35	0.00022	PRK05379	PRK05379	cl28366	bifunctional nicotinamide mononucleotide adenylyltransferase
	superfamily	4	35	0.00022	cl28366	PRK05379 superfamily	-	bifunctional nicotinamide mononucleotide adenylyltransferase
	non-specific	3	46	0.000235	TIGR01527	arch_NMN_Atrans	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	1	69	0.000255	COG1019	CAB4	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	7	67	0.001261	cd02164	PPAT_CoAS	cl00015	phosphopantetheine adenylyltransferase
	non-specific	13	80	0.001583	PLN02660	PLN02660	cl00015	pantoate--beta-alanine ligase

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. lipofaciens	non-specific	13	39	0.002191	COG0414	PanC	cl00015	Panthothenate synthetase
	non-specific	5	72	0.002316	TIGR01518	g3p_cytidyltrns	cl00015	glycerol-3-phosphate cytidylyltransferase
	non-specific	13	39	0.003169	cd00560	PanC	cl00015	Pantoate-beta-alanine ligase
	non-specific	13	39	0.005229	PRK00380	panC	cl00015	pantoate--beta-alanine ligase
M. mobile	non-specific	9	96	1.63e-14	PRK13964	coaD	cl00015	phosphopantetheine adenylyltransferase
	superfamily	9	96	1.63e-14	cl00015	nt_trans superfamily	-	nucleotidyl transferase superfamily
	non-specific	11	96	8.75e-14	TIGR01510	coaD_prev_kdtB	cl00015	pantetheine-phosphate adenylyltransferase
	non-specific	8	96	1.02e-12	COG0669	CoaD	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	11	96	3.11e-12	cd02163	PPAT	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	8	96	1.73e-11	PRK00168	coaD	cl00015	phosphopantetheine adenylyltransferase
	non-specific	11	58	2.49e-05	TIGR00125	cyt_tran_rel	cl00015	cytidyltransferase-like domain
M. molare	non-specific	1	140	2.91e-62	PRK13964	coaD	cl00015	phosphopantetheine adenylyltransferase
	superfamily	1	140	2.91e-62	cl00015	nt_trans superfamily	-	nucleotidyl transferase superfamily
	non-specific	1	140	3.14e-45	COG0669	CoaD	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	3	131	5.35e-42	cd02163	PPAT	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	1	131	2.39e-40	PRK00168	coaD	cl00015	phosphopantetheine adenylyltransferase
	non-specific	3	133	2.95e-38	TIGR01510	coaD_prev_kdtB	cl00015	pantetheine-phosphate adenylyltransferase
	specific	5	135	8.78e-22	pfam01467	CTP_transf_like	cl00015	Cytidylyltransferase-like
	non-specific	3	63	7.16e-12	TIGR00125	cyt_tran_rel	cl00015	cytidyltransferase-like domain
	non-specific	3	133	1.09e-09	cd02039	cytidylyltransferase_like	cl00015	Cytidylyltransferase-like domain
	non-specific	1	44	7.06e-07	PRK00071	nadD	cl00015	nicotinic acid mononucleotide adenylyltransferase
	non-specific	3	90	1.32e-06	cd02156	nt_trans	cl00015	nucleotidyl transferase superfamily
	non-specific	1	40	1.97e-06	COG1056	NadR	cl00015	Nicotinamide mononucleotide adenylyltransferase
	non-specific	4	66	2.69e-06	COG1057	NadD	cl00015	Nicotinic acid mononucleotide adenylyltransferase
	non-specific	4	66	4.66e-06	cd02165	NMNAT	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	1	66	8.56e-06	PRK07152	nadD	cl28367	putative nicotinate-nucleotide adenylyltransferase
	superfamily	1	66	8.56e-06	cl28367	nadD superfamily	-	putative nicotinate-nucleotide adenylyltransferase
	non-specific	3	35	1.19e-05	PRK01153	PRK01153	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	3	35	2.27e-05	cd02166	NMNAT_Archaea	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	5	66	5.12e-05	TIGR00482	TIGR00482	cl00015	nicotinate (nicotinamide) nucleotide adenylyltransferase
	non-specific	7	41	0.000146	cd02173	ECT	cl00015	CTP:phosphoethanolamine cytidylyltransferase (ECT)
	non-specific	1	36	0.00028	PRK00777	PRK00777	cl00015	phosphopantetheine adenylyltransferase
	non-specific	4	35	0.000284	PRK05379	PRK05379	cl28366	bifunctional nicotinamide mononucleotide adenylyltransferase
	superfamily	4	35	0.000284	cl28366	PRK05379 superfamily	-	bifunctional nicotinamide mononucleotide adenylyltransferase
	non-specific	12	67	0.000349	smart00764	Citrate_ly_lig	cl00015	Citrate lyase ligase C-terminal domain
	non-specific	12	66	0.000385	pfam08218	Citrate_ly_lig	cl00015	Citrate lyase ligase C-terminal domain
	non-specific	8	44	0.000398	TIGR02199	rfaE_dom_II	cl00015	rfaE bifunctional protein, domain II
	non-specific	1	85	0.000414	COG0615	TagD	cl00015	Glycerol-3-phosphate cytidylyltransferase, cytidylyltransferase family

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. molare	non-specific	12	71	0.000594	cd02169	Citrate_lyase_ligase	cl00015	Citrate lyase ligase
	non-specific	1	57	0.001273	cd02170	cytidylyltransferase	cl00015	cytidylyltransferase
	non-specific	8	41	0.001836	cd02174	CCT	cl00015	CTP:phosphocholine cytidylyltransferase
	non-specific	7	41	0.001934	PTZ00308	PTZ00308	cl28626	ethanolamine-phosphate cytidylyltransferase
	superfamily	7	41	0.001934	cl28626	PLN02406 superfamily	-	ethanolamine-phosphate cytidylyltransferase
	non-specific	8	133	0.002042	COG2870	RfaE	cl28454	ADP-heptose synthase, bifunctional sugar kinase/adenylyltransferase
	superfamily	8	133	0.002042	cl28454	RfaE superfamily	-	ADP-heptose synthase, bifunctional sugar kinase/adenylyltransferase
	non-specific	11	64	0.002251	COG3053	CitC	cl28578	Citrate lyase synthetase [Energy production and conversion]
	superfamily	11	64	0.002251	cl28578	CitC superfamily	-	Citrate lyase synthetase [Energy production and conversion]
	non-specific	1	55	0.005139	cd02171	G3P_Cytidylyltransferase	cl00015	glycerol-3-phosphate cytidylyltransferase
M. mycoides subsp. *capri*	specific	1	140	2.94e-71	PRK13964	coaD	cl00015	phosphopantetheine adenylyltransferase
	superfamily	1	140	2.94e-71	cl00015	nt_trans superfamily	-	nucleotidyl transferase superfamily
	non-specific	3	140	1.91e-62	TIGR01510	coaD_prev_kdtB	cl00015	pantetheine-phosphate adenylyltransferase
	non-specific	1	139	1.17e-45	COG0669	CoaD	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	1	135	6.52e-44	PRK00168	coaD	cl00015	phosphopantetheine adenylyltransferase
	non-specific	3	139	1.02e-42	cd02163	PPAT	cl00015	Phosphopantetheine adenylyltransferase
	specific	5	135	7.19e-20	pfam01467	CTP_transf_like	cl00015	Cytidylyltransferase-like
	non-specific	3	63	2.98e-13	TIGR00125	cyt_tran_rel	cl00015	cytidyltransferase-like domain
	non-specific	3	134	8.76e-12	cd02039	cytidylyltransferase_like	cl00015	Cytidylyltransferase-like domain
	non-specific	1	77	6.01e-10	PRK07152	nadD	cl28367	putative nicotinate-nucleotide adenylyltransferase
	superfamily	1	77	6.01e-10	cl28367	nadD superfamily	-	putative nicotinate-nucleotide adenylyltransferase
	non-specific	1	66	1.52e-08	COG1057	NadD	cl00015	Nicotinic acid mononucleotide adenylyltransferase
	non-specific	1	136	4.04e-08	PRK00777	PRK00777	cl00015	phosphopantetheine adenylyltransferase
	non-specific	3	77	1.56e-07	cd02165	NMNAT	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	1	44	2.58e-07	COG1056	NadR	cl00015	Nicotinamide mononucleotide adenylyltransferase
	non-specific	1	136	6.67e-07	COG1019	CAB4	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	1	66	1.60e-06	PRK00071	nadD	cl00015	nicotinic acid mononucleotide adenylyltransferase
	non-specific	1	71	3.91e-05	TIGR01526	nadR_NMN_Atrans	cl28365	nicotinamide-nucleotide adenylyltransferase
	superfamily	1	71	3.91e-05	cl28365	PRK08099 superfamily	-	bifunctional DNA-binding transcriptional repressor/ NMN adenylyltransferase
	non-specific	4	61	5.48e-05	cd02166	NMNAT_Archaea	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	3	134	6.30e-05	cd02167	NMNAT_NadR	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	4	130	6.91e-05	cd02156	nt_trans	cl00015	nucleotidyl transferase superfamily
	non-specific	4	68	7.26e-05	TIGR01527	arch_NMN_Atrans	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	3	35	0.00049	PRK05379	PRK05379	cl28366	bifunctional nicotinamide mononucleotide adenylyltransferase

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. mycoides subsp. *capri*	superfamily	3	35	0.00049	cl28366	PRK05379 superfamily	-	bifunctional nicotinamide mononucleotide adenylyltransferase
	non-specific	4	35	0.000534	PRK01153	PRK01153	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	1	76	0.000596	PRK08887	PRK08887	cl00015	nicotinic acid mononucleotide adenylyltransferase
	non-specific	8	89	0.001492	cd02170	cytidylyltransferase	cl00015	cytidylyltransferase
	non-specific	2	67	0.001944	PRK01170	PRK01170	cl00866	phosphopantetheine adenylyltransferase
	superfamily	2	67	0.001944	cl00866	NTPase_I-T superfamily	-	Protein of unknown function DUF84
	non-specific	8	133	0.002054	cd02164	PPAT_CoAS	cl00015	phosphopantetheine adenylyltransferase
	non-specific	1	36	0.003852	cd02171	G3P_Cytidylyltransferase	cl00015	glycerol-3-phosphate cytidylyltransferase
	non-specific	5	36	0.003895	TIGR00482	TIGR00482	cl00015	nicotinate (nicotinamide) nucleotide adenylyltransferase
M. mycoides subsp. *mycoides*	specific	1	140	4.77e-72	PRK13964	coaD	cl00015	phosphopantetheine adenylyltransferase
	superfamily	1	140	4.77e-72	cl00015	nt_trans superfamily	-	nucleotidyl transferase superfamily
	non-specific	3	140	8.20e-62	TIGR01510	coaD_prev_kdtB	cl00015	pantetheine-phosphate adenylyltransferase
	non-specific	1	139	1.19e-45	COG0669	CoaD	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	1	135	3.07e-43	PRK00168	coaD	cl00015	phosphopantetheine adenylyltransferase
	non-specific	3	139	1.96e-42	cd02163	PPAT	cl00015	Phosphopantetheine adenylyltransferase
	specific	5	135	1.38e-20	pfam01467	CTP_transf_like	cl00015	Cytidylyltransferase-like
	non-specific	4	63	8.39e-13	TIGR00125	cyt_tran_rel	cl00015	cytidyltransferase-like domain
	non-specific	4	134	2.80e-12	cd02039	cytidylyltransferase_like	cl00015	Cytidylyltransferase-like domain
	non-specific	1	124	2.63e-10	PRK07152	nadD	cl28367	putative nicotinate-nucleotide adenylyltransferase
	superfamily	1	124	2.63e-10	cl28367	nadD superfamily	-	putative nicotinate-nucleotide adenylyltransferase
	non-specific	1	66	7.45e-08	COG1057	NadD	cl00015	Nicotinic acid mononucleotide adenylyltransferase
	non-specific	1	66	1.14e-06	PRK00071	nadD	cl00015	nicotinic acid mononucleotide adenylyltransferase
	non-specific	3	77	3.95e-06	cd02165	NMNAT	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	3	134	4.26e-06	cd02167	NMNAT_NadR	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	1	136	8.25e-06	PRK00777	PRK00777	cl00015	phosphopantetheine adenylyltransferase
	non-specific	1	35	1.04e-05	COG1056	NadR	cl00015	Nicotinamide mononucleotide adenylyltransferase
	non-specific	3	130	1.73e-05	cd02156	nt_trans	cl00015	nucleotidyl transferase superfamily
	non-specific	1	71	0.000106	TIGR01526	nadR_NMN_Atrans	cl28365	nicotinamide-nucleotide adenylyltransferase
	superfamily	1	71	0.000106	cl28365	PRK08099 superfamily	-	bifunctional DNA-binding transcriptional repressor/ NMN adenylyltransferase
	non-specific	5	119	0.000126	TIGR00482	TIGR00482	cl00015	nicotinate (nicotinamide) nucleotide adenylyltransferase
	non-specific	4	35	0.000356	PRK01153	PRK01153	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	1	136	0.000358	COG1019	CAB4	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	4	35	0.000496	cd02166	NMNAT_Archaea	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	4	35	0.000519	PRK05379	PRK05379	cl28366	bifunctional nicotinamide mononucleotide adenylyltransferase
	superfamily	4	35	0.000519	cl28366	PRK05379 superfamily	-	bifunctional nicotinamide mononucleotide adenylyltransferase
	non-specific	1	36	0.002311	cd02171	G3P_Cytidylyltransferase	cl00015	glycerol-3-phosphate cytidylyltransferase

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. mycoides subsp. *mycoides*	non-specific	1	36	0.003558	cd02170	cytidylyltransferase	cl00015	cytidylyltransferase
	non-specific	10	35	0.006714	pfam05636	HIGH_NTase1	cl27012	HIGH Nucleotidyl Transferase
	superfamily	10	35	0.006714	cl27012	HIGH_NTase1 superfamily	-	HIGH Nucleotidyl Transferase
	non-specific	4	68	0.007739	TIGR01527	arch_NMN_Atrans	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	10	35	0.007951	PRK13670	PRK13670	cl27012	hypothetical protein
M. opalescens	non-specific	1	141	1.49e-55	PRK13964	coaD	cl00015	phosphopantetheine adenylyltransferase
	superfamily	1	141	1.49e-55	cl00015	nt_trans superfamily	-	nucleotidyl transferase superfamily
	non-specific	1	135	7.10e-39	COG0669	CoaD	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	3	134	7.30e-35	TIGR01510	coaD_prev_kdtB	cl00015	pantetheine-phosphate adenylyltransferase
	non-specific	3	132	3.82e-34	cd02163	PPAT	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	1	132	2.96e-32	PRK00168	coaD	cl00015	phosphopantetheine adenylyltransferase
	non-specific	5	136	2.16e-15	pfam01467	CTP_transf_like	cl00015	Cytidylyltransferase-like
	non-specific	3	51	2.96e-10	TIGR00125	cyt_tran_rel	cl00015	cytidyltransferase-like domain
	non-specific	1	43	7.33e-09	COG1057	NadD	cl00015	Nicotinic acid mononucleotide adenylyltransferase
	non-specific	1	67	2.07e-08	PRK07152	nadD	cl28367	putative nicotinate-nucleotide adenylyltransferase
	superfamily	1	67	2.07e-08	cl28367	nadD superfamily	-	putative nicotinate-nucleotide adenylyltransferase
	non-specific	5	67	3.93e-08	TIGR00482	TIGR00482	cl00015	nicotinate (nicotinamide) nucleotide adenylyltransferase
	non-specific	5	134	2.86e-07	cd02167	NMNAT_NadR	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	4	67	4.08e-07	cd02165	NMNAT	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	1	35	3.27e-06	COG1056	NadR	cl00015	Nicotinamide mononucleotide adenylyltransferase
	non-specific	1	44	4.36e-06	PRK00071	nadD	cl00015	nicotinic acid mononucleotide adenylyltransferase
	non-specific	3	134	5.17e-06	cd02039	cytidylyltransferase_like	cl00015	Cytidylyltransferase-like domain
	non-specific	3	39	5.97e-06	cd02156	nt_trans	cl00015	nucleotidyl transferase superfamily
	non-specific	1	61	6.36e-06	cd02171	G3P_Cytidylyltransferase	cl00015	glycerol-3-phosphate cytidylyltransferase
	non-specific	1	88	2.33e-05	cd02170	cytidylyltransferase	cl00015	cytidylyltransferase
	non-specific	1	36	3.02e-05	COG0615	TagD	cl00015	Glycerol-3-phosphate cytidylyltransferase, cytidylyltransferase family
	non-specific	1	33	0.000173	PRK00777	PRK00777	cl00015	phosphopantetheine adenylyltransferase
	non-specific	1	37	0.000471	TIGR01526	nadR_NMN_Atrans	cl28365	nicotinamide-nucleotide adenylyltransferase
	superfamily	1	37	0.000471	cl28365	PRK08099 superfamily	-	bifunctional DNA-binding transcriptional repressor/ NMN adenylyltransferase
	non-specific	5	61	0.001081	TIGR01518	g3p_cytidyltrns	cl00015	glycerol-3-phosphate cytidylyltransferase
	non-specific	14	42	0.001136	pfam02569	Pantoate_ligase	cl00015	Pantoate-beta-alanine ligase
	non-specific	2	29	0.001861	PRK01153	PRK01153	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	8	50	0.002073	TIGR02199	rfaE_dom_II	cl00015	rfaE bifunctional protein, domain II
	non-specific	4	35	0.00227	PRK05379	PRK05379	cl28366	bifunctional nicotinamide mononucleotide adenylyltransferase
	superfamily	4	35	0.00227	cl28366	PRK05379 superfamily	-	bifunctional nicotinamide mononucleotide adenylyltransferase
	non-specific	1	42	0.002519	PRK08099	PRK08099	cl28365	bifunctional DNA-binding transcriptional repressor/ NMN adenylyltransferase

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. opalescens	non-specific	3	35	0.004583	cd02166	NMNAT_Archaea	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
M. penetrans	non-specific	10	136	8.33e-36	TIGR01510	coaD_prev_kdtB	cl00015	pantetheine-phosphate adenylyltransferase
	superfamily	10	136	8.33e-36	cl00015	nt_trans superfamily	-	nucleotidyl transferase superfamily
	non-specific	6	135	1.92e-31	PRK00168	coaD	cl00015	phosphopantetheine adenylyltransferase
	non-specific	7	135	1.58e-30	COG0669	CoaD	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	11	136	4.53e-29	cd02163	PPAT	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	7	146	3.67e-25	PRK13964	coaD	cl00015	phosphopantetheine adenylyltransferase
	non-specific	11	139	2.00e-11	pfam01467	CTP_transf_like	cl00015	Cytidylyltransferase-like
	non-specific	14	140	2.37e-08	cd02170	cytidylyltransferase	cl00015	cytidylyltransferase
	non-specific	11	42	3.71e-08	TIGR00125	cyt_tran_rel	cl00015	cytidyltransferase-like domain
	non-specific	14	140	1.07e-06	COG0615	TagD	cl00015	Glycerol-3-phosphate cytidylyltransferase, cytidylyltransferase family
	non-specific	10	141	4.00e-06	cd02039	cytidylyltransferase_like	cl00015	Cytidylyltransferase-like domain
	non-specific	5	41	1.80e-05	COG1056	NadR	cl00015	Nicotinamide mononucleotide adenylyltransferase
	non-specific	11	95	3.02e-05	cd02156	nt_trans	cl00015	nucleotidyl transferase superfamily
	non-specific	14	76	4.52e-05	PRK00777	PRK00777	cl00015	phosphopantetheine adenylyltransferase
	non-specific	14	89	0.000143	COG1019	CAB4	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	8	93	0.000157	cd02171	G3P_Cytidylyltransferase	cl00015	glycerol-3-phosphate cytidylyltransferase
	non-specific	12	41	0.001478	cd02166	NMNAT_Archaea	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	9	41	0.00178	PRK05379	PRK05379	cl28366	bifunctional nicotinamide mononucleotide adenylyltransferase
	superfamily	9	41	0.00178	cl28366	PRK05379 superfamily	-	bifunctional nicotinamide mononucleotide adenylyltransferase
	non-specific	6	73	0.003098	PRK00071	nadD	cl00015	nicotinic acid mononucleotide adenylyltransferase
	non-specific	14	45	0.004532	cd02174	CCT	cl00015	CTP:phosphocholine cytidylyltransferase
	non-specific	10	93	0.004773	cd02165	NMNAT	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	2	83	0.005485	PRK07143	PRK07143	cl27514	hypothetical protein
	superfamily	2	83	0.005485	cl27514	Flavokinase superfamily	-	Riboflavin kinase
	non-specific	13	50	0.005531	cd02173	ECT	cl00015	CTP:phosphoethanolamine cytidylyltransferase (ECT)
	non-specific	6	68	0.005561	COG1057	NadD	cl00015	Nicotinic acid mononucleotide adenylyltransferase
	non-specific	11	41	0.006622	PRK01153	PRK01153	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	14	93	0.006986	TIGR01518	g3p_cytidyltrns	cl00015	glycerol-3-phosphate cytidylyltransferase
M. pirum	non-specific	5	136	6.06e-34	cd02163	PPAT	cl00015	Phosphopantetheine adenylyltransferase
	superfamily	5	136	6.06e-34	cl00015	nt_trans superfamily	-	nucleotidyl transferase superfamily
	non-specific	3	136	1.16e-33	COG0669	CoaD	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	5	136	6.94e-30	TIGR01510	coaD_prev_kdtB	cl00015	pantetheine-phosphate adenylyltransferase
	non-specific	3	136	2.01e-29	PRK00168	coaD	cl00015	phosphopantetheine adenylyltransferase
	non-specific	3	148	1.60e-27	PRK13964	coaD	cl00015	phosphopantetheine adenylyltransferase
	non-specific	7	137	1.78e-09	pfam01467	CTP_transf_like	cl00015	Cytidylyltransferase-like
	non-specific	5	136	1.17e-05	cd02039	cytidylyltransferase_like	cl00015	Cytidylyltransferase-like domain

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. pirum	non-specific	5	64	5.58e-05	TIGR00125	cyt_tran_rel	cl00015	cytidyltransferase-like domain
	non-specific	1	139	0.003566	COG1057	NadD	cl00015	Nicotinic acid mononucleotide adenylyltransferase
M. primatum	specific	1	140	4.60e-68	PRK13964	coaD	cl00015	phosphopantetheine adenylyltransferase
	superfamily	1	140	4.60e-68	cl00015	nt_trans superfamily	-	nucleotidyl transferase superfamily
	non-specific	1	139	7.16e-45	COG0669	CoaD	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	3	133	4.10e-40	cd02163	PPAT	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	3	133	1.77e-39	TIGR01510	coaD_prev_kdtB	cl00015	pantetheine-phosphate adenylyltransferase
	non-specific	1	133	6.48e-39	PRK00168	coaD	cl00015	phosphopantetheine adenylyltransferase
	specific	5	135	2.92e-19	pfam01467	CTP_transf_like	cl00015	Cytidylyltransferase-like
	non-specific	3	63	1.25e-11	TIGR00125	cyt_tran_rel	cl00015	cytidyltransferase-like domain
	non-specific	4	66	5.00e-10	cd02165	NMNAT	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	1	66	2.27e-09	PRK07152	nadD	cl28367	putative nicotinate-nucleotide adenylyltransferase
	superfamily	1	66	2.27e-09	cl28367	nadD superfamily	-	putative nicotinate-nucleotide adenylyltransferase
	non-specific	3	133	3.18e-09	cd02039	cytidylyltransferase_like	cl00015	Cytidylyltransferase-like domain
	non-specific	5	133	2.80e-08	cd02167	NMNAT_NadR	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	1	66	2.88e-08	COG1057	NadD	cl00015	Nicotinic acid mononucleotide adenylyltransferase
	non-specific	1	66	3.34e-07	PRK00071	nadD	cl00015	nicotinic acid mononucleotide adenylyltransferase
	non-specific	1	87	4.45e-07	COG0615	TagD	cl00015	Glycerol-3-phosphate cytidylyltransferase, cytidylyltransferase family
	non-specific	1	35	4.80e-07	COG1056	NadR	cl00015	Nicotinamide mononucleotide adenylyltransferase
	non-specific	1	36	5.86e-07	cd02170	cytidylyltransferase	cl00015	cytidylyltransferase
	non-specific	5	56	1.62e-06	TIGR00482	TIGR00482	cl00015	nicotinate (nicotinamide) nucleotide adenylyltransferase
	non-specific	3	39	1.08e-05	cd02156	nt_trans	cl00015	nucleotidyl transferase superfamily
	non-specific	1	36	5.76e-05	cd02171	G3P_Cytidylyltransferase	cl00015	glycerol-3-phosphate cytidylyltransferase
	non-specific	5	58	6.21e-05	TIGR01518	g3p_cytidyltrns	cl00015	glycerol-3-phosphate cytidylyltransferase
	non-specific	1	134	9.28e-05	PRK08099	PRK08099	cl28365	bifunctional DNA-binding transcriptional repressor/ NMN adenylyltransferase
	superfamily	1	134	9.28e-05	cl28365	PRK08099 superfamily	-	bifunctional DNA-binding transcriptional repressor/ NMN adenylyltransferase
	non-specific	13	39	0.000104	pfam02569	Pantoate_ligase	cl00015	Pantoate-beta-alanine ligase
	non-specific	1	33	0.000108	PRK00777	PRK00777	cl00015	phosphopantetheine adenylyltransferase
	non-specific	1	133	0.000184	TIGR01526	nadR_NMN_Atrans	cl28365	nicotinamide-nucleotide adenylyltransferase
	non-specific	2	35	0.000231	PRK01153	PRK01153	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	3	42	0.00048	cd02173	ECT	cl00015	CTP:phosphoethanolamine cytidylyltransferase (ECT)
	non-specific	3	35	0.000531	cd02166	NMNAT_Archaea	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	1	89	0.00074	COG1019	CAB4	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	3	46	0.001026	TIGR01527	arch_NMN_Atrans	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	13	39	0.001088	cd00560	PanC	cl00015	Pantoate-beta-alanine ligase
	non-specific	4	35	0.001237	PRK05379	PRK05379	cl28366	bifunctional nicotinamide mononucleotide adenylyltransferase

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. primatum	superfamily	4	35	0.001237	cl28366	PRK05379 superfamily	-	bifunctional nicotinamide mononucleotide adenylyltransferase
	non-specific	13	39	0.001776	COG0414	PanC	cl00015	Panthothenate synthetase
	non-specific	13	39	0.00428	PRK00380	panC	cl00015	pantoate--beta-alanine ligase
	non-specific	7	60	0.005073	cd02164	PPAT_CoAS	cl00015	phosphopantetheine adenylyltransferase
M. pulmonis	specific	5	146	9.05e-69	PRK13964	coaD	cl00015	phosphopantetheine adenylyltransferase
	superfamily	5	146	9.05e-69	cl00015	nt_trans superfamily	-	nucleotidyl transferase superfamily
	specific	4	146	8.15e-57	COG0669	CoaD	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	7	139	1.27e-39	cd02163	PPAT	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	5	139	5.99e-38	PRK00168	coaD	cl00015	phosphopantetheine adenylyltransferase
	non-specific	7	139	7.37e-38	TIGR01510	coaD_prev_kdtB	cl00015	pantetheine-phosphate adenylyltransferase
	specific	9	139	1.04e-19	pfam01467	CTP_transf_like	cl00015	Cytidylyltransferase-like
	non-specific	2	59	4.15e-12	PRK00071	nadD	cl00015	nicotinic acid mononucleotide adenylyltransferase
	non-specific	7	67	2.35e-11	TIGR00125	cyt_tran_rel	cl00015	cytidyltransferase-like domain
	non-specific	3	59	1.26e-09	COG1057	NadD	cl00015	Nicotinic acid mononucleotide adenylyltransferase
	non-specific	7	59	2.23e-09	cd02165	NMNAT	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	7	139	6.78e-09	cd02039	cytidylyltransferase_like	cl00015	Cytidylyltransferase-like domain
	non-specific	5	70	6.06e-07	PRK07152	nadD	cl28367	putative nicotinate-nucleotide adenylyltransferase
	superfamily	5	70	6.06e-07	cl28367	nadD superfamily	-	putative nicotinate-nucleotide adenylyltransferase
	non-specific	9	93	2.24e-06	TIGR00482	TIGR00482	cl00015	nicotinate (nicotinamide) nucleotide adenylyltransferase
	non-specific	1	70	4.89e-06	PRK05379	PRK05379	cl28366	bifunctional nicotinamide mononucleotide adenylyltransferase
	superfamily	1	70	4.89e-06	cl28366	PRK05379 superfamily	-	bifunctional nicotinamide mononucleotide adenylyltransferase
	non-specific	8	63	8.69e-06	cd02156	nt_trans	cl00015	nucleotidyl transferase superfamily
	non-specific	3	39	0.000122	COG1056	NadR	cl00015	Nicotinamide mononucleotide adenylyltransferase
	non-specific	7	139	0.000176	cd02064	FAD_synthetase_N	cl00015	FAD synthetase, N-terminal domain of the bifunctional enzyme
	non-specific	5	71	0.000571	PRK08887	PRK08887	cl00015	nicotinic acid mononucleotide adenylyltransferase
	non-specific	8	39	0.000958	PRK01153	PRK01153	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	4	46	0.001423	PTZ00308	PTZ00308	cl28626	ethanolamine-phosphate cytidylyltransferase
	superfamily	4	46	0.001423	cl28626	PLN02406 superfamily	-	ethanolamine-phosphate cytidylyltransferase
	non-specific	12	47	0.001912	cd02174	CCT	cl00015	CTP:phosphocholine cytidylyltransferase
	non-specific	7	146	0.00194	cd02167	NMNAT_NadR	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	5	139	0.003724	PRK00777	PRK00777	cl00015	phosphopantetheine adenylyltransferase
	non-specific	8	39	0.004703	cd02166	NMNAT_Archaea	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	7	29	0.004914	cd02168	NMNAT_Nudix	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	6	67	0.005299	PRK01170	PRK01170	cl00866	phosphopantetheine adenylyltransferase
	superfamily	6	67	0.005299	cl00866	NTPase_I-T superfamily	-	Protein of unknown function DUF84

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. pulmonis	non-specific	1	139	0.005711	COG1019	CAB4	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	6	60	0.007006	cd02171	G3P_Cytidylyltransferase	cl00015	glycerol-3-phosphate cytidylyltransferase
M. putrefaciens	non-specific	3	141	1.92e-54	TIGR01510	coaD_prev_kdtB	cl00015	pantetheine-phosphate adenylyltransferase
	superfamily	3	141	1.92e-54	cl00015	nt_trans superfamily	-	nucleotidyl transferase superfamily
	non-specific	1	140	1.67e-53	PRK13964	coaD	cl00015	phosphopantetheine adenylyltransferase
	non-specific	1	141	1.78e-53	PRK00168	coaD	cl00015	phosphopantetheine adenylyltransferase
	non-specific	3	141	1.79e-50	cd02163	PPAT	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	1	141	1.96e-50	COG0669	CoaD	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	5	135	4.78e-15	pfam01467	CTP_transf_like	cl00015	Cytidylyltransferase-like
	non-specific	3	134	2.73e-12	cd02039	cytidylyltransferase_like	cl00015	Cytidylyltransferase-like domain
	non-specific	3	63	4.33e-12	TIGR00125	cyt_tran_rel	cl00015	cytidyltransferase-like domain
	non-specific	1	66	1.60e-09	COG1057	NadD	cl00015	Nicotinic acid mononucleotide adenylyltransferase
	non-specific	1	66	4.28e-08	PRK07152	nadD	cl28367	putative nicotinate-nucleotide adenylyltransferase
	superfamily	1	66	4.28e-08	cl28367	nadD superfamily	-	putative nicotinate-nucleotide adenylyltransferase
	non-specific	3	66	2.30e-07	cd02165	NMNAT	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	1	35	2.44e-07	COG1056	NadR	cl00015	Nicotinamide mononucleotide adenylyltransferase
	non-specific	1	66	1.14e-06	PRK00071	nadD	cl00015	nicotinic acid mononucleotide adenylyltransferase
	non-specific	1	64	1.61e-05	TIGR01526	nadR_NMN_Atrans	cl28365	nicotinamide-nucleotide adenylyltransferase
	superfamily	1	64	1.61e-05	cl28365	PRK08099 superfamily	-	bifunctional DNA-binding transcriptional repressor/ NMN adenylyltransferase
	non-specific	8	134	1.66e-05	COG0615	TagD	cl00015	Glycerol-3-phosphate cytidylyltransferase, cytidylyltransferase family
	non-specific	4	66	2.10e-05	TIGR01527	arch_NMN_Atrans	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	3	39	4.29e-05	PRK05379	PRK05379	cl28366	bifunctional nicotinamide mononucleotide adenylyltransferase
	superfamily	3	39	4.29e-05	cl28366	PRK05379 superfamily	-	bifunctional nicotinamide mononucleotide adenylyltransferase
	non-specific	4	35	0.00012	cd02166	NMNAT_Archaea	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	8	97	0.000166	cd02170	cytidylyltransferase	cl00015	cytidylyltransferase
	non-specific	1	135	0.000191	cd02171	G3P_Cytidylyltransferase	cl00015	glycerol-3-phosphate cytidylyltransferase
	non-specific	4	35	0.000215	PRK01153	PRK01153	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	3	66	0.000497	cd02167	NMNAT_NadR	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	3	24	0.000885	cd02168	NMNAT_Nudix	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	4	130	0.001521	cd02156	nt_trans	cl00015	nucleotidyl transferase superfamily
	non-specific	5	85	0.002241	TIGR00482	TIGR00482	cl00015	nicotinate (nicotinamide) nucleotide adenylyltransferase
	non-specific	1	33	0.002402	PRK00777	PRK00777	cl00015	phosphopantetheine adenylyltransferase
	non-specific	8	35	0.002999	cd02174	CCT	cl00015	CTP:phosphocholine cytidylyltransferase
M. simbae	non-specific	1	140	7.69e-56	PRK13964	coaD	cl00015	phosphopantetheine adenylyltransferase
	superfamily	1	140	7.69e-56	cl00015	nt_trans superfamily	-	nucleotidyl transferase superfamily

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. simbae	non-specific	1	134	2.68e-43	COG0669	CoaD	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	3	133	6.47e-40	cd02163	PPAT	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	1	133	1.02e-38	PRK00168	coaD	cl00015	phosphopantetheine adenylyltransferase
	non-specific	3	133	2.41e-35	TIGR01510	coaD_prev_kdtB	cl00015	pantetheine-phosphate adenylyltransferase
	specific	5	135	3.01e-18	pfam01467	CTP_transf_like	cl00015	Cytidylyltransferase-like
	non-specific	5	133	1.67e-09	cd02167	NMNAT_NadR	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	3	46	1.73e-09	TIGR00125	cyt_tran_rel	cl00015	cytidyltransferase-like domain
	non-specific	3	133	1.59e-07	cd02039	cytidylyltransferase_like	cl00015	Cytidylyltransferase-like domain
	non-specific	1	35	1.63e-07	COG1056	NadR	cl00015	Nicotinamide mononucleotide adenylyltransferase
	non-specific	1	66	2.64e-07	COG1057	NadD	cl00015	Nicotinic acid mononucleotide adenylyltransferase
	non-specific	1	66	2.74e-07	PRK07152	nadD	cl28367	putative nicotinate-nucleotide adenylyltransferase
	superfamily	1	66	2.74e-07	cl28367	nadD superfamily	-	putative nicotinate-nucleotide adenylyltransferase
	non-specific	4	41	5.16e-07	cd02165	NMNAT	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	1	52	4.38e-06	PRK00071	nadD	cl00015	nicotinic acid mononucleotide adenylyltransferase
	non-specific	5	66	7.52e-06	TIGR00482	TIGR00482	cl00015	nicotinate (nicotinamide) nucleotide adenylyltransferase
	non-specific	1	133	1.05e-05	TIGR01526	nadR_NMN_Atrans	cl28365	nicotinamide-nucleotide adenylyltransferase
	superfamily	1	133	1.05e-05	cl28365	PRK08099 superfamily	-	bifunctional DNA-binding transcriptional repressor/ NMN adenylyltransferase
	non-specific	1	33	1.95e-05	PRK00777	PRK00777	cl00015	phosphopantetheine adenylyltransferase
	non-specific	4	68	2.47e-05	PRK05379	PRK05379	cl28366	bifunctional nicotinamide mononucleotide adenylyltransferase
	superfamily	4	68	2.47e-05	cl28366	PRK05379 superfamily	-	bifunctional nicotinamide mononucleotide adenylyltransferase
	non-specific	3	35	3.94e-05	cd02166	NMNAT_Archaea	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	1	35	5.06e-05	PRK01153	PRK01153	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	3	56	7.86e-05	TIGR01527	arch_NMN_Atrans	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	8	134	0.000172	PRK08099	PRK08099	cl28365	bifunctional DNA-binding transcriptional repressor/ NMN adenylyltransferase
	non-specific	4	42	0.000416	cd02168	NMNAT_Nudix	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	1	56	0.000768	COG1019	CAB4	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	1	36	0.001184	cd02171	G3P_Cytidylyltransferase	cl00015	glycerol-3-phosphate cytidylyltransferase
	non-specific	8	27	0.002505	pfam06574	FAD_syn	cl00015	FAD synthetase
	non-specific	7	59	0.004095	cd02164	PPAT_CoAS	cl00015	phosphopantetheine adenylyltransferase
	non-specific	3	57	0.00418	cd02156	nt_trans	cl00015	nucleotidyl transferase superfamily
	non-specific	1	43	0.004248	COG0615	TagD	cl00015	Glycerol-3-phosphate cytidylyltransferase, cytidylyltransferase family
	non-specific	8	50	0.004811	PTZ00308	PTZ00308	cl28626	ethanolamine-phosphate cytidylyltransferase
	superfamily	8	50	0.004811	cl28626	PLN02406 superfamily	-	ethanolamine-phosphate cytidylyltransferase
M. sturni	non-specific	3	142	5.48e-59	PRK13964	coaD	cl00015	phosphopantetheine adenylyltransferase
	superfamily	3	142	5.48e-59	cl00015	nt_trans superfamily	-	nucleotidyl transferase superfamily

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. sturni	non-specific	4	144	8.41e-51	COG0669	CoaD	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	5	144	4.45e-46	cd02163	PPAT	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	4	144	2.23e-43	PRK00168	coaD	cl00015	phosphopantetheine adenylyltransferase
	non-specific	5	144	3.64e-42	TIGR01510	coaD_prev_kdtB	cl00015	pantetheine-phosphate adenylyltransferase
	specific	7	135	7.56e-27	pfam01467	CTP_transf_like	cl00015	Cytidylyltransferase-like
	non-specific	5	65	6.76e-14	TIGR00125	cyt_tran_rel	cl00015	cytidyltransferase-like domain
	non-specific	1	68	2.33e-12	PRK00071	nadD	cl00015	nicotinic acid mononucleotide adenylyltransferase
	non-specific	1	68	2.90e-11	COG1057	NadD	cl00015	Nicotinic acid mononucleotide adenylyltransferase
	non-specific	4	68	1.19e-10	PRK07152	nadD	cl28367	putative nicotinate-nucleotide adenylyltransferase
	superfamily	4	68	1.19e-10	cl28367	nadD superfamily	-	putative nicotinate-nucleotide adenylyltransferase
	non-specific	5	135	3.63e-10	cd02039	cytidylyltransferase_like	cl00015	Cytidylyltransferase-like domain
	non-specific	5	68	4.80e-10	cd02165	NMNAT	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	1	136	6.18e-08	PRK08099	PRK08099	cl28365	bifunctional DNA-binding transcriptional repressor/ NMN adenylyltransferase
	superfamily	1	136	6.18e-08	cl28365	PRK08099 superfamily	-	bifunctional DNA-binding transcriptional repressor/ NMN adenylyltransferase
	non-specific	7	64	2.74e-07	TIGR00482	TIGR00482	cl00015	nicotinate (nicotinamide) nucleotide adenylyltransferase
	non-specific	4	135	6.18e-07	TIGR01526	nadR_NMN_Atrans	cl28365	nicotinamide-nucleotide adenylyltransferase
	non-specific	1	37	6.44e-07	COG1056	NadR	cl00015	Nicotinamide mononucleotide adenylyltransferase
	non-specific	5	68	2.02e-06	cd02167	NMNAT_NadR	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	6	92	2.35e-06	cd02156	nt_trans	cl00015	nucleotidyl transferase superfamily
	non-specific	10	72	3.65e-05	PRK00777	PRK00777	cl00015	phosphopantetheine adenylyltransferase
	non-specific	6	37	8.73e-05	PRK01153	PRK01153	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	6	37	0.000192	cd02166	NMNAT_Archaea	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	7	37	0.000354	cd02173	ECT	cl00015	CTP:phosphoethanolamine cytidylyltransferase (ECT)
	non-specific	10	78	0.00064	COG1019	CAB4	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	4	37	0.000661	PTZ00308	PTZ00308	cl28626	ethanolamine-phosphate cytidylyltransferase
	superfamily	4	37	0.000661	cl28626	PLN02406 superfamily	-	ethanolamine-phosphate cytidylyltransferase
	non-specific	4	43	0.000784	TIGR02199	rfaE_dom_II	cl00015	rfaE bifunctional protein, domain II
	non-specific	10	38	0.001375	cd02170	cytidylyltransferase	cl00015	cytidylyltransferase
	non-specific	10	42	0.003324	COG0615	TagD	cl00015	Glycerol-3-phosphate cytidylyltransferase, cytidylyltransferase family
	non-specific	1	36	0.006569	PRK05379	PRK05379	cl28366	bifunctional nicotinamide mononucleotide adenylyltransferase
	superfamily	1	36	0.006569	cl28366	PRK05379 superfamily	-	bifunctional nicotinamide mononucleotide adenylyltransferase
M. synoviae	specific	7	145	1.02e-70	PRK13964	coaD	cl00015	phosphopantetheine adenylyltransferase
	superfamily	7	145	1.02e-70	cl00015	nt_trans superfamily	-	nucleotidyl transferase superfamily
	non-specific	5	139	8.48e-46	COG0669	CoaD	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	8	138	4.70e-43	cd02163	PPAT	cl00015	Phosphopantetheine adenylyltransferase

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. synoviae	non-specific	7	138	7.70e-39	PRK00168	coaD	cl00015	phosphopantetheine adenylyltransferase
	non-specific	8	138	2.29e-38	TIGR01510	coaD_prev_kdtB	cl00015	pantetheine-phosphate adenylyltransferase
	specific	10	138	1.75e-21	pfam01467	CTP_transf_like	cl00015	Cytidylyltransferase-like
	non-specific	8	40	3.49e-12	TIGR00125	cyt_tran_rel	cl00015	cytidyltransferase-like domain
	non-specific	8	138	2.88e-11	cd02039	cytidylyltransferase_like	cl00015	Cytidylyltransferase-like domain
	non-specific	7	71	5.01e-08	COG1057	NadD	cl00015	Nicotinic acid mononucleotide adenylyltransferase
	non-specific	8	44	6.67e-08	cd02156	nt_trans	cl00015	nucleotidyl transferase superfamily
	non-specific	10	138	7.22e-08	cd02167	NMNAT_NadR	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	9	71	7.22e-08	cd02165	NMNAT	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	3	46	1.03e-06	PRK00071	nadD	cl00015	nicotinic acid mononucleotide adenylyltransferase
	non-specific	13	41	1.37e-06	cd02170	cytidylyltransferase	cl00015	cytidylyltransferase
	non-specific	10	67	1.89e-06	TIGR00482	TIGR00482	cl00015	nicotinate (nicotinamide) nucleotide adenylyltransferase
	non-specific	13	90	2.20e-06	COG0615	TagD	cl00015	Glycerol-3-phosphate cytidylyltransferase, cytidylyltransferase family
	non-specific	7	48	6.80e-06	PRK07152	nadD	cl28367	putative nicotinate-nucleotide adenylyltransferase
	superfamily	7	48	6.80e-06	cl28367	nadD superfamily	-	putative nicotinate-nucleotide adenylyltransferase
	non-specific	7	40	2.13e-05	COG1056	NadR	cl00015	Nicotinamide mononucleotide adenylyltransferase
	non-specific	13	38	0.000116	PRK00777	PRK00777	cl00015	phosphopantetheine adenylyltransferase
	non-specific	10	41	0.000464	TIGR01526	nadR_NMN_Atrans	cl28365	nicotinamide-nucleotide adenylyltransferase
	superfamily	10	41	0.000464	cl28365	PRK08099 superfamily	-	bifunctional DNA-binding transcriptional repressor/ NMN adenylyltransferase
	non-specific	13	46	0.000556	cd02174	CCT	cl00015	CTP:phosphocholine cytidylyltransferase
	non-specific	7	40	0.001029	PRK01153	PRK01153	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	8	40	0.001086	cd02166	NMNAT_Archaea	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	12	46	0.001356	PTZ00308	PTZ00308	cl28626	ethanolamine-phosphate cytidylyltransferase
	superfamily	12	46	0.001356	cl28626	PLN02406 superfamily	-	ethanolamine-phosphate cytidylyltransferase
	non-specific	7	41	0.001661	cd02171	G3P_Cytidylyltransferase	cl00015	glycerol-3-phosphate cytidylyltransferase
	non-specific	8	46	0.002191	cd02173	ECT	cl00015	CTP:phosphoethanolamine cytidylyltransferase (ECT)
	non-specific	13	62	0.002334	TIGR01518	g3p_cytidyltrns	cl00015	glycerol-3-phosphate cytidylyltransferase
	non-specific	3	88	0.007463	COG1019	CAB4	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	13	139	0.009413	PRK08099	PRK08099	cl28365	bifunctional DNA-binding transcriptional repressor/ NMN adenylyltransferase
M. testudinis	non-specific	4	138	6.38e-51	cd02163	PPAT	cl00015	Phosphopantetheine adenylyltransferase
	superfamily	4	138	6.38e-51	cl00015	nt_trans superfamily	-	nucleotidyl transferase superfamily
	non-specific	3	138	5.73e-50	PRK00168	coaD	cl00015	phosphopantetheine adenylyltransferase
	non-specific	1	138	6.30e-50	COG0669	CoaD	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	4	138	4.10e-46	TIGR01510	coaD_prev_kdtB	cl00015	pantetheine-phosphate adenylyltransferase
	non-specific	3	145	6.67e-38	PRK13964	coaD	cl00015	phosphopantetheine adenylyltransferase
	specific	6	137	1.45e-21	pfam01467	CTP_transf_like	cl00015	Cytidylyltransferase-like
	non-specific	4	62	5.82e-13	TIGR00125	cyt_tran_rel	cl00015	cytidyltransferase-like domain

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. testudinis	non-specific	1	87	9.30e-13	PRK00071	nadD	cl00015	nicotinic acid mononucleotide adenylyltransferase
	non-specific	4	135	5.30e-11	cd02039	cytidylyltransferase_like	cl00015	Cytidylyltransferase-like domain
	non-specific	4	142	7.57e-11	cd02165	NMNAT	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	1	136	3.00e-10	COG1057	NadD	cl00015	Nicotinic acid mononucleotide adenylyltransferase
	non-specific	5	36	9.11e-09	cd02166	NMNAT_Archaea	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	5	36	2.08e-08	PRK01153	PRK01153	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	3	36	6.77e-08	COG1056	NadR	cl00015	Nicotinamide mononucleotide adenylyltransferase
	non-specific	1	142	1.49e-07	PRK08887	PRK08887	cl00015	nicotinic acid mononucleotide adenylyltransferase
	non-specific	6	103	8.88e-07	TIGR00482	TIGR00482	cl00015	nicotinate (nicotinamide) nucleotide adenylyltransferase
	non-specific	5	48	2.28e-06	cd02156	nt_trans	cl00015	nucleotidyl transferase superfamily
	non-specific	5	67	2.29e-05	TIGR01527	arch_NMN_Atrans	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	3	87	2.33e-05	PRK07152	nadD	cl28367	putative nicotinate-nucleotide adenylyltransferase
	superfamily	3	87	2.33e-05	cl28367	nadD superfamily	-	putative nicotinate-nucleotide adenylyltransferase
	non-specific	4	33	5.56e-05	PRK05379	PRK05379	cl28366	bifunctional nicotinamide mononucleotide adenylyltransferase
	superfamily	4	33	5.56e-05	cl28366	PRK05379 superfamily	-	bifunctional nicotinamide mononucleotide adenylyltransferase
	non-specific	2	36	8.10e-05	PTZ00308	PTZ00308	cl28626	ethanolamine-phosphate cytidylyltransferase
	superfamily	2	36	8.10e-05	cl28626	PLN02406 superfamily	-	ethanolamine-phosphate cytidylyltransferase
	non-specific	9	42	8.99e-05	COG0615	TagD	cl00015	Glycerol-3-phosphate cytidylyltransferase, cytidylyltransferase family
	non-specific	9	56	9.03e-05	cd02170	cytidylyltransferase	cl00015	cytidylyltransferase
	non-specific	3	36	0.000179	cd02173	ECT	cl00015	CTP:phosphoethanolamine cytidylyltransferase (ECT)
	non-specific	4	36	0.000766	cd02168	NMNAT_Nudix	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	3	37	0.001214	cd02171	G3P_Cytidylyltransferase	cl00015	glycerol-3-phosphate cytidylyltransferase
	non-specific	1	36	0.003957	PRK00777	PRK00777	cl00015	phosphopantetheine adenylyltransferase
M. yeatsii	non-specific	3	139	2.57e-56	TIGR01510	coaD_prev_kdtB	cl00015	pantetheine-phosphate adenylyltransferase
	superfamily	3	139	2.57e-56	cl00015	nt_trans superfamily	-	nucleotidyl transferase superfamily
	non-specific	1	135	9.05e-55	PRK00168	coaD	cl00015	phosphopantetheine adenylyltransferase
	non-specific	1	140	1.17e-53	PRK13964	coaD	cl00015	phosphopantetheine adenylyltransferase
	non-specific	1	140	1.38e-52	COG0669	CoaD	cl00015	Phosphopantetheine adenylyltransferase
	non-specific	3	139	1.53e-52	cd02163	PPAT	cl00015	Phosphopantetheine adenylyltransferase
	specific	5	135	3.84e-20	pfam01467	CTP_transf_like	cl00015	Cytidylyltransferase-like
	non-specific	3	63	1.68e-14	TIGR00125	cyt_tran_rel	cl00015	cytidyltransferase-like domain
	non-specific	3	134	1.00e-13	cd02039	cytidylyltransferase_like	cl00015	Cytidylyltransferase-like domain
	non-specific	3	66	1.63e-09	cd02165	NMNAT	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	1	35	5.73e-09	COG1056	NadR	cl00015	Nicotinamide mononucleotide adenylyltransferase
	non-specific	1	60	2.21e-08	COG1057	NadD	cl00015	Nicotinic acid mononucleotide adenylyltransferase
	non-specific	1	66	1.18e-07	PRK07152	nadD	cl28367	putative nicotinate-nucleotide adenylyltransferase

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. yeatsii	superfamily	1	66	1.18e-07	cl28367	nadD superfamily	-	putative nicotinate-nucleotide adenylyltransferase
	non-specific	4	66	1.83e-07	TIGR01527	arch_NMN_Atrans	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	1	57	2.01e-07	cd02171	G3P_Cytidylyltransferase	cl00015	glycerol-3-phosphate cytidylyltransferase
	non-specific	4	35	3.01e-07	PRK01153	PRK01153	cl00015	nicotinamide-nucleotide adenylyltransferase
	non-specific	4	35	5.09e-07	cd02166	NMNAT_Archaea	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	8	89	5.98e-07	cd02170	cytidylyltransferase	cl00015	cytidylyltransferase
	non-specific	1	69	7.93e-07	TIGR01526	nadR_NMN_Atrans	cl28365	nicotinamide-nucleotide adenylyltransferase
	superfamily	1	69	7.93e-07	cl28365	PRK08099 superfamily	-	bifunctional DNA-binding transcriptional repressor/ NMN adenylyltransferase
	non-specific	3	55	9.84e-07	PRK05379	PRK05379	cl28366	bifunctional nicotinamide mononucleotide adenylyltransferase
	superfamily	3	55	9.84e-07	cl28366	PRK05379 superfamily	-	bifunctional nicotinamide mononucleotide adenylyltransferase
	non-specific	1	66	1.10e-06	PRK00071	nadD	cl00015	nicotinic acid mononucleotide adenylyltransferase
	non-specific	4	130	4.04e-06	cd02156	nt_trans	cl00015	nucleotidyl transferase superfamily
	non-specific	3	66	1.23e-05	cd02167	NMNAT_NadR	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	8	138	4.20e-05	COG0615	TagD	cl00015	Glycerol-3-phosphate cytidylyltransferase, cytidylyltransferase family
	non-specific	3	76	5.40e-05	cd02168	NMNAT_Nudix	cl00015	Nicotinamide/nicotinate mononucleotide adenylyltransferase
	non-specific	1	33	0.000427	PRK00777	PRK00777	cl00015	phosphopantetheine adenylyltransferase
	non-specific	5	54	0.001052	TIGR00482	TIGR00482	cl00015	nicotinate (nicotinamide) nucleotide adenylyltransferase
	non-specific	8	54	0.001145	TIGR01518	g3p_cytidyltrns	cl00015	glycerol-3-phosphate cytidylyltransferase
	non-specific	8	134	0.001474	PRK07143	PRK07143	cl27514	hypothetical protein
	superfamily	8	134	0.001474	cl27514	Flavokinase superfamily	-	Riboflavin kinase
	non-specific	2	35	0.001999	PTZ00308	PTZ00308	cl28626	ethanolamine-phosphate cytidylyltransferase
	superfamily	2	35	0.001999	cl28626	PLN02406 superfamily	-	ethanolamine-phosphate cytidylyltransferase
	non-specific	8	134	0.002964	COG0196	RibF	cl27514	FAD synthase
	non-specific	5	33	0.006077	cd02173	ECT	cl00015	CTP:phosphoethanolamine cytidylyltransferase (ECT)
	non-specific	12	73	0.00806	cd02169	Citrate_lyase_ligase	cl00015	Citrate lyase ligase

Supplementary Table 1.6 The CDD search results of the DPCK amino acid sequences of the currently annotated CoA biosynthetic pathway enzyme-encoding genes and newly identified homologues

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. sp. Ms02	non-specific	1	122	1.50e-13	COG0237	CoaE	cl28605	Dephospho-CoA kinase
	superfamily	1	122	1.50e-13	cl28605	CoaE superfamily	-	Dephospho-CoA kinase
	non-specific	1	122	6.33e-11	cd02022	DPCK	cl17190	Dephospho-CoA kinase
	superfamily	1	122	6.33e-11	cl17190	NK superfamily	-	Nucleoside/nucleotide kinase superfamily
	non-specific	2	122	2.20e-08	TIGR00152	TIGR00152	cl17190	Dephospho-CoA kinase
	non-specific	2	120	2.87e-05	PRK00081	coaE	cl17190	Dephospho-CoA kinase
	non-specific	2	48	0.000637	cd01983	Fer4_NifH	cl28886	The Fer4_NifH superfamily
	superfamily	2	48	0.000637	cl28886	Fer4_NifH superfamily	-	The Fer4_NifH superfamily
	non-specific	1	33	0.001471	COG1936	Fap7	cl17190	Broad-specificity NMP kinase
M. agalactiae	non-specific	1	143	2.34e-23	COG0237	CoaE	cl28605	Dephospho-CoA kinase
	superfamily	1	143	2.34e-23	cl28605	CoaE superfamily	-	Dephospho-CoA kinase
	specific	1	143	8.86e-21	cd02022	DPCK	cl17190	Dephospho-CoA kinase
	superfamily	1	143	8.86e-21	cl17190	NK superfamily	-	Nucleoside/nucleotide kinase superfamily
	non-specific	2	186	2.85e-14	TIGR00152	TIGR00152	cl17190	Dephospho-CoA kinase
	non-specific	2	146	2.73e-13	PRK00081	coaE	cl17190	Dephospho-CoA kinase
	non-specific	2	142	1.59e-06	pfam01121	CoaE	cl17190	Dephospho-CoA kinase
	non-specific	1	121	9.07e-06	COG1936	Fap7	cl17190	Broad-specificity NMP kinase
	non-specific	1	92	0.004474	PLN02422	PLN02422	cl17190	Dephospho-CoA kinase
	non-specific	2	129	0.007723	COG0572	Udk	cl28606	Uridine kinase superfamily
	superfamily	2	129	0.007723	cl28606	Udk superfamily	-	Uridine kinase superfamily
	non-specific	2	35	0.009588	PRK03839	PRK03839	cl17190	putative kinase
M. alligatoris	non-specific	1	132	1.34e-11	COG0237	CoaE	cl28605	Dephospho-CoA kinase
	superfamily	1	132	1.34e-11	cl28605	CoaE superfamily	-	Dephospho-CoA kinase
	non-specific	1	132	4.77e-11	cd02022	DPCK	cl17190	Dephospho-CoA kinase
	superfamily	1	132	4.77e-11	cl17190	NK superfamily	-	Nucleoside/nucleotide kinase superfamily
	non-specific	2	132	2.79e-05	PRK00081	coaE	cl17190	Dephospho-CoA kinase
	non-specific	2	136	0.000429	TIGR00152	TIGR00152	cl17190	Dephospho-CoA kinase
M. alvi	non-specific	2	188	2.29e-24	TIGR00152	TIGR00152	cl17190	Dephospho-CoA kinase
	superfamily	2	188	2.29e-24	cl17190	NK superfamily	-	Nucleoside/nucleotide kinase superfamily
	specific	2	151	1.45e-21	cd02022	DPCK	cl17190	Dephospho-CoA kinase
	non-specific	1	191	1.66e-20	PRK00081	coaE	cl17190	Dephospho-CoA kinase
	non-specific	1	191	5.20e-20	COG0237	CoaE	cl28605	Dephospho-CoA kinase
	superfamily	1	191	5.20e-20	cl28605	CoaE superfamily	-	Dephospho-CoA kinase
	non-specific	1	99	2.52e-12	pfam01121	CoaE	cl17190	Dephospho-CoA kinase
	non-specific	2	106	8.98e-11	PLN02422	PLN02422	cl17190	Dephospho-CoA kinase
	non-specific	5	89	1.98e-08	PRK14730	coaE	cl17190	Dephospho-CoA kinase
	non-specific	2	194	2.83e-06	PRK14732	coaE	cl17190	Dephospho-CoA kinase
	non-specific	1	21	0.002608	PRK04040	PRK04040	cl17190	adenylate kinase

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. alvi	non-specific	2	44	0.003877	cd03255	ABC_MJ0796_LolCDE_FtsE	cl25403	ATP-binding cassette domain of the transporters involved in export of lipoprotein and macrolide, and cell division protein
	superfamily	2	44	0.003877	cl25403	ABC_ATPase superfamily	-	ATP-binding cassette transporter nucleotide-binding domain
	non-specific	1	21	0.005459	COG2019	AdkA	cl17190	Archaeal adenylate kinase
	non-specific	1	21	0.005635	COG1936	Fap7	cl17190	Broad-specificity NMP kinase
	non-specific	2	44	0.00671	cd03263	ABC_subfamily_A	cl25403	ATP-binding cassette domain of the lipid transporters, subfamily A
	non-specific	2	50	0.007481	TIGR02982	heterocyst_DevA	cl28181	ABC exporter ATP-binding subunit, DevA family
	superfamily	2	50	0.007481	cl28181	AAA superfamily	-	ATPases associated with a variety of cellular activities
M. anatis	non-specific	1	137	2.61e-13	COG0237	CoaE	cl28605	Dephospho-CoA kinase
	superfamily	1	137	2.61e-13	cl28605	CoaE superfamily	-	Dephospho-CoA kinase
	non-specific	1	137	2.26e-09	cd02022	DPCK	cl17190	Dephospho-CoA kinase
	superfamily	1	137	2.26e-09	cl17190	NK superfamily	-	Nucleoside/nucleotide kinase superfamily
	non-specific	2	137	4.69e-05	TIGR00152	TIGR00152	cl17190	Dephospho-CoA kinase
	non-specific	2	137	4.71e-05	PRK00081	coaE	cl17190	Dephospho-CoA kinase
	non-specific	2	133	0.000588	pfam01121	CoaE	cl17190	Dephospho-CoA kinase
	non-specific	2	190	0.001564	PRK14733	coaE	cl17190	Dephospho-CoA kinase
M. arginini	non-specific	1	141	5.25e-13	COG0237	CoaE	cl28605	Dephospho-CoA kinase
	superfamily	1	141	5.25e-13	cl28605	CoaE superfamily	-	Dephospho-CoA kinase
	non-specific	1	147	2.86e-11	cd02022	DPCK	cl17190	Dephospho-CoA kinase
	superfamily	1	147	2.86e-11	cl17190	NK superfamily	-	Nucleoside/nucleotide kinase superfamily
	non-specific	2	166	5.29e-07	TIGR00152	TIGR00152	cl17190	Dephospho-CoA kinase
	non-specific	2	147	0.006534	PRK00081	coaE	cl17190	Dephospho-CoA kinase
	non-specific	1	76	0.009556	pfam05272	VirE	cl23993	Virulence-associated protein E (these proteins contain a P-loop motif)
	superfamily	1	76	0.009556	cl23993	VirE superfamily	-	Virulence-associated protein E (these proteins contain a P-loop motif)
M. bovigenitalium	non-specific	1	190	2.24e-17	COG0237	CoaE	cl28605	Dephospho-CoA kinase
	superfamily	1	190	2.24e-17	cl28605	CoaE superfamily	-	Dephospho-CoA kinase
	non-specific	1	178	3.64e-12	cd02022	DPCK	cl17190	Dephospho-CoA kinase
	superfamily	1	178	3.64e-12	cl17190	NK superfamily	-	Nucleoside/nucleotide kinase superfamily
	non-specific	2	137	2.67e-07	TIGR00152	TIGR00152	cl17190	Dephospho-CoA kinase
	non-specific	2	88	6.80e-06	PRK00081	coaE	cl17190	Dephospho-CoA kinase
	non-specific	1	117	0.00034	PRK14732	coaE	cl17190	Dephospho-CoA kinase
M. bovis	non-specific	1	149	4.53e-21	COG0237	CoaE	cl28605	Dephospho-CoA kinase
	superfamily	1	149	4.53e-21	cl28605	CoaE superfamily	-	Dephospho-CoA kinase
	specific	1	146	6.23e-20	cd02022	DPCK	cl17190	Dephospho-CoA kinase
	superfamily	1	146	6.23e-20	cl17190	NK superfamily	-	Nucleoside/nucleotide kinase superfamily
	non-specific	2	177	1.36e-11	TIGR00152	TIGR00152	cl17190	Dephospho-CoA kinase
	non-specific	2	146	4.16e-11	PRK00081	coaE	cl17190	Dephospho-CoA kinase

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. bovis	non-specific	1	78	2.40e-06	COG1936	Fap7	cl17190	Broad-specificity NMP kinase
	non-specific	1	91	0.00063	PRK14731	coaE	cl17190	Dephospho-CoA kinase
	non-specific	2	33	0.001113	cd01983	Fer4_NifH	cl28886	The Fer4_NifH superfamily
	superfamily	2	33	0.001113	cl28886	Fer4_NifH superfamily	-	The Fer4_NifH superfamily
	non-specific	2	148	0.001637	pfam13238	AAA_18	cl21455	AAA domain
	superfamily	2	148	0.001637	cl21455	P-loop_NTPase superfamily	-	P-loop containing Nucleoside Triphosphate Hydrolases
M. bovoculi (HAD-DPCK)	specific	1	263	9.47e-37	COG0561	Cof	cl26787	Hydroxymethylpyrimidine pyrophosphatase and other HAD family phosphatases [Coenzyme transport and metabolism, General function prediction only]
	superfamily	1	263	9.47e-37	cl26787	Hydrolase_3 superfamily	-	haloacid dehalogenase-like hydrolase
	non-specific	7	242	7.44e-36	pfam08282	Hydrolase_3	cl26787	haloacid dehalogenase-like hydrolase
	non-specific	7	256	2.02e-33	TIGR00099	Cof-subfamily	cl26787	Cof subfamily of IIB subfamily of haloacid dehalogenase superfamily
	non-specific	7	258	7.41e-31	cd07516	HAD_Pase	cl21460	phosphatase, similar to Escherichia coli Cof and Thermotoga maritima TM0651
	superfamily	7	258	7.41e-31	cl21460	HAD_like superfamily	-	Haloacid Dehalogenase-like Hydrolases
	non-specific	1	261	1.23e-14	PRK01158	PRK01158	cl26787	phosphoglycolate phosphatase
	non-specific	4	244	6.09e-14	cd07517	HAD_HPP	cl21460	phosphatase, similar to Bacteroides thetaiotaomicron VPI-5482 BT4131 hexose phosphate phosphatase
	non-specific	8	259	3.25e-12	TIGR01482	SPP-subfamily	cl26787	sucrose-phosphate phosphatase subfamily
	non-specific	266	400	5.02e-12	COG0237	CoaE	cl28605	Dephospho-CoA kinase
	superfamily	266	400	5.02e-12	cl28605	CoaE superfamily	-	Dephospho-CoA kinase
	non-specific	266	393	2.03e-11	cd02022	DPCK	cl17190	Dephospho-CoA kinase
	superfamily	266	393	2.03e-11	cl17190	NK superfamily	-	Nucleoside/nucleotide kinase superfamily
	non-specific	7	218	1.49e-10	TIGR01484	HAD-SF-IIB	cl26787	HAD-superfamily
	non-specific	9	218	2.48e-09	cd02605	HAD_SPP	cl21460	sucrose-phosphatase, similar to Synechocystis sp PCC 6803 SPP
	non-specific	4	244	4.82e-08	TIGR01487	Pglycolate_arch	cl26787	phosphoglycolate phosphatase, TA0175-type
	non-specific	266	391	9.33e-07	TIGR00152	TIGR00152	cl17190	Dephospho-CoA kinase
	non-specific	3	240	9.70e-07	PRK10976	PRK10976	cl26787	putative hydrolase
	non-specific	1	261	1.95e-06	PRK10513	PRK10513	cl26787	sugar phosphate phosphatase
	non-specific	266	394	2.91e-05	PRK00081	coaE	cl17190	Dephospho-CoA kinase
	non-specific	148	257	3.67e-05	cd07518	HAD_YbiV-Like	cl21460	Escherichia coli YbiV sugar phosphatase/phosphotransferase and related proteins
	non-specific	2	259	7.30e-05	PLN02887	PLN02887	cl26787	hydrolase family protein
	non-specific	5	75	0.00369	TIGR01486	HAD-SF-IIB-MPGP	cl26786	mannosyl-3-phosphoglycerate phosphatase family
	superfamily	5	75	0.00369	cl26786	YedP superfamily	-	Predicted mannosyl-3-phosphoglycerate phosphatase, HAD superfamily
	non-specific	178	218	0.004772	pfam05116	S6PP	cl26787	Sucrose-6F-phosphate phosphohydrolase

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. bovoculi (HAD-DPCK)	non-specific	175	259	0.008491	cd07514	HAD_Pase	cl21460	phosphatase, similar to Thermoplasma acidophilum TA0175 phosphoglycolate phosphatase (PCPase), and Pyrococcus horikoshii PH1421, a magnesium-dependent phosphatase
M. buteonis	non-specific	1	185	5.06e-12	COG0237	CoaE	cl28605	Dephospho-CoA kinase
	superfamily	1	185	5.06e-12	cl28605	CoaE superfamily	-	Dephospho-CoA kinase
	non-specific	1	129	2.49e-09	cd02022	DPCK	cl17190	Dephospho-CoA kinase
	superfamily	1	129	2.49e-09	cl17190	NK superfamily	-	Nucleoside/nucleotide kinase superfamily
	non-specific	2	128	2.01e-06	TIGR00152	TIGR00152	cl17190	Dephospho-CoA kinase
M. californicum	non-specific	1	189	3.53e-15	COG0237	CoaE	cl28605	Dephospho-CoA kinase
	superfamily	1	189	3.53e-15	cl28605	CoaE superfamily	-	Dephospho-CoA kinase
	non-specific	1	148	1.98e-14	cd02022	DPCK	cl17190	Dephospho-CoA kinase
	superfamily	1	148	1.98e-14	cl17190	NK superfamily	-	Nucleoside/nucleotide kinase superfamily
	non-specific	2	148	5.93e-08	PRK00081	coaE	cl17190	Dephospho-CoA kinase
	non-specific	2	189	6.62e-08	TIGR00152	TIGR00152	cl17190	Dephospho-CoA kinase
	non-specific	2	137	0.000139	pfam01121	CoaE	cl17190	Dephospho-CoA kinase
M. canis	non-specific	1	129	8.82e-10	COG0237	CoaE	cl28605	Dephospho-CoA kinase
	superfamily	1	129	8.82e-10	cl28605	CoaE superfamily	-	Dephospho-CoA kinase
	non-specific	2	173	5.03e-06	TIGR00152	TIGR00152	cl17190	Dephospho-CoA kinase
	superfamily	2	173	5.03e-06	cl17190	NK superfamily	-	Nucleoside/nucleotide kinase superfamily
	non-specific	1	129	1.37e-05	cd02022	DPCK	cl17190	Dephospho-CoA kinase
	non-specific	2	34	0.001443	cd01983	Fer4_NifH	cl28886	The Fer4_NifH superfamily
	superfamily	2	34	0.001443	cl28886	Fer4_NifH superfamily	-	The Fer4_NifH superfamily
	non-specific	1	88	0.006613	COG0488	Uup	cl28181	ATPase components of ABC transporters with duplicated ATPase domains [General function prediction only]
	superfamily	1	88	0.006613	cl28181	AAA superfamily	-	ATPases associated with a variety of cellular activities
M. capricolum subsp. *capricolum*	specific	6	179	2.24e-38	cd02022	DPCK	cl17190	Dephospho-CoA kinase
	superfamily	6	179	2.24e-38	cl17190	NK superfamily	-	Nucleoside/nucleotide kinase superfamily
	non-specific	4	174	1.85e-26	PRK00081	coaE	cl17190	Dephospho-CoA kinase
	non-specific	4	179	1.75e-23	COG0237	CoaE	cl28605	Dephospho-CoA kinase
	superfamily	4	179	1.75e-23	cl28605	CoaE superfamily	-	Dephospho-CoA kinase
	non-specific	5	174	3.75e-23	pfam01121	CoaE	cl17190	Dephospho-CoA kinase
	non-specific	6	174	7.13e-18	TIGR00152	TIGR00152	cl17190	Dephospho-CoA kinase
	non-specific	6	188	5.99e-12	PRK14730	coaE	cl17190	Dephospho-CoA kinase
	non-specific	6	179	7.13e-11	PRK14732	coaE	cl17190	Dephospho-CoA kinase
	non-specific	7	174	4.72e-08	PRK14733	coaE	cl17190	Dephospho-CoA kinase
	non-specific	6	179	2.23e-05	PLN02422	PLN02422	cl17190	Dephospho-CoA kinase
	non-specific	7	174	0.000153	PRK14734	coaE	cl17190	Dephospho-CoA kinase
	non-specific	3	174	0.000277	PRK14731	coaE	cl17190	Dephospho-CoA kinase

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. capricolum subsp. *capripneumoniae*	specific	2	175	1.30e-36	cd02022	DPCK	cl17190	Dephospho-CoA kinase
	superfamily	2	175	1.30e-36	cl17190	NK superfamily	-	Nucleoside/nucleotide kinase superfamily
	non-specific	1	170	1.66e-24	PRK00081	coaE	cl17190	Dephospho-CoA kinase
	non-specific	1	175	1.93e-22	COG0237	CoaE	cl28605	Dephospho-CoA kinase
	superfamily	1	175	1.93e-22	cl28605	CoaE superfamily	-	Dephospho-CoA kinase
	non-specific	1	170	3.53e-22	pfam01121	CoaE	cl17190	Dephospho-CoA kinase
	non-specific	2	170	1.19e-16	TIGR00152	TIGR00152	cl17190	Dephospho-CoA kinase
	non-specific	2	184	2.12e-11	PRK14730	coaE	cl17190	Dephospho-CoA kinase
	non-specific	2	175	2.21e-11	PRK14732	coaE	cl17190	Dephospho-CoA kinase
	non-specific	3	170	5.81e-08	PRK14733	coaE	cl17190	Dephospho-CoA kinase
	non-specific	2	175	0.001038	PLN02422	PLN02422	cl17190	Dephospho-CoA kinase
	non-specific	2	170	0.001838	PRK14731	coaE	cl17190	Dephospho-CoA kinase
M. collis	specific	1	129	9.11e-25	COG0237	CoaE	cl28605	Dephospho-CoA kinase
	superfamily	1	129	9.11e-25	cl28605	CoaE superfamily	-	Dephospho-CoA kinase
	specific	1	148	1.10e-20	cd02022	DPCK	cl17190	Dephospho-CoA kinase
	superfamily	1	148	1.10e-20	cl17190	NK superfamily	-	Nucleoside/nucleotide kinase superfamily
	non-specific	2	148	3.80e-17	TIGR00152	TIGR00152	cl17190	Dephospho-CoA kinase
	non-specific	2	148	6.58e-15	PRK00081	coaE	cl17190	Dephospho-CoA kinase
	non-specific	2	150	4.58e-07	PRK14730	coaE	cl17190	Dephospho-CoA kinase
	non-specific	1	80	4.87e-07	COG1936	Fap7	cl17190	Broad-specificity NMP kinase
	non-specific	2	129	9.87e-06	pfam01121	CoaE	cl17190	Dephospho-CoA kinase
	non-specific	2	33	2.26e-05	cd01983	Fer4_NifH	cl28886	The Fer4_NifH superfamily
	superfamily	2	33	2.26e-05	cl28886	Fer4_NifH superfamily	-	The Fer4_NifH superfamily
	non-specific	2	80	0.000107	PRK03839	PRK03839	cl17190	putative kinase
	non-specific	2	92	0.000163	PRK04182	PRK04182	cl28332	cytidylate kinase
	superfamily	2	92	0.000163	cl28332	CmkB superfamily	-	Cytidylate kinase
	non-specific	1	86	0.000597	PLN02422	PLN02422	cl17190	Dephospho-CoA kinase
	specific	2	124	0.000634	smart00382	AAA	cl28181	ATPases associated with a variety of cellular activities
	superfamily	2	124	0.000634	cl28181	AAA superfamily	-	ATPases associated with a variety of cellular activities
	non-specific	2	73	0.001659	COG1102	CmkB	cl28332	Cytidylate kinase
	non-specific	2	42	0.001875	PRK04040	PRK04040	cl17190	adenylate kinase
	non-specific	1	20	0.002192	cd03255	ABC_MJ0796_LolCDE_FtsE	cl25403	ATP-binding cassette domain of the transporters involved in export of lipoprotein and macrolide, and cell division protein
	superfamily	1	20	0.002192	cl25403	ABC_ATPase superfamily	-	ATP-binding cassette transporter nucleotide-binding domain
	non-specific	5	81	0.00335	pfam13207	AAA_17	cl21455	AAA domain
	superfamily	5	81	0.00335	cl21455	P-loop_NTPase superfamily	-	P-loop containing Nucleoside Triphosphate Hydrolases

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. collis	non-specific	1	26	0.003507	cd03116	MobB	cl28886	Molybdenum is an essential trace element in the form of molybdenum cofactor (Moco) which is associated with the metabolism of nitrogen, carbon and sulfur by redox active enzymes. In E. coli, the synthesis of Moco involves genes from several loci: moa, mob, mod, moe and mog. The mob locus contains mobA and mobB genes. MobB catalyzes the attachment of the guanine dinucleotide to molybdopterin.
	non-specific	2	26	0.003758	pfam13521	AAA_28	cl21455	AAA domain
	non-specific	2	145	0.00394	PRK14733	coaE	cl17190	Dephospho-CoA kinase
	non-specific	1	20	0.006373	COG1136	LolD	cl28181	ABC-type lipoprotein export system, ATPase component [Cell wall/membrane/envelope biogenesis]
	non-specific	2	73	0.00661	cd02034	CooC	cl28886	The accessory protein CooC, which contains a nucleotide-binding domain (P-loop) near the N-terminus, participates in the maturation of the nickel center of carbon monoxide dehydrogenase (CODH).
	non-specific	2	151	0.007951	cd01851	GBP	cl21455	Guanylate-binding protein (GBP) family (N-terminal domain)
	non-specific	1	92	0.009239	TIGR02173	cyt_kin_arch	cl28332	cytidylate kinase
	non-specific	2	28	0.009861	cd02020	CMPK	cl17190	Cytidine monophosphate kinase
M. columbinum	non-specific	1	137	1.47e-22	COG0237	CoaE	cl28605	Dephospho-CoA kinase
	superfamily	1	137	1.47e-22	cl28605	CoaE superfamily	-	Dephospho-CoA kinase
	specific	1	137	4.33e-22	cd02022	DPCK	cl17190	Dephospho-CoA kinase
	superfamily	1	137	4.33e-22	cl17190	NK superfamily	-	Nucleoside/nucleotide kinase superfamily
	non-specific	2	137	1.79e-12	PRK00081	coaE	cl17190	Dephospho-CoA kinase
	non-specific	2	140	5.95e-12	TIGR00152	TIGR00152	cl17190	Dephospho-CoA kinase
	non-specific	2	137	2.38e-07	pfam01121	CoaE	cl17190	Dephospho-CoA kinase
	non-specific	1	88	0.000169	PRK14732	coaE	cl17190	Dephospho-CoA kinase
	non-specific	2	126	0.000305	pfam13238	AAA_18	cl21455	AAA domain
	superfamily	2	126	0.000305	cl21455	P-loop_NTPase superfamily	-	P-loop containing Nucleoside Triphosphate Hydrolases
	non-specific	1	89	0.000568	PRK14731	coaE	cl17190	Dephospho-CoA kinase
	non-specific	9	136	0.007099	pfam13207	AAA_17	cl21455	AAA domain
M. columborale	non-specific	1	181	4.64e-11	COG0237	CoaE	cl28605	Dephospho-CoA kinase
	superfamily	1	181	4.64e-11	cl28605	CoaE superfamily	-	Dephospho-CoA kinase
	non-specific	1	144	1.93e-07	cd02022	DPCK	cl17190	Dephospho-CoA kinase
	superfamily	1	144	1.93e-07	cl17190	NK superfamily	-	Nucleoside/nucleotide kinase superfamily
	non-specific	1	156	2.71e-05	PRK00081	coaE	cl17190	Dephospho-CoA kinase
	non-specific	2	136	0.000544	TIGR00152	TIGR00152	cl17190	Dephospho-CoA kinase
	non-specific	4	137	0.000723	PRK14731	coaE	cl17190	Dephospho-CoA kinase
	non-specific	1	93	0.000804	PRK14732	coaE	cl17190	Dephospho-CoA kinase

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. columborale	non-specific	1	30	0.006918	PRK14734	coaE	cl17190	Dephospho-CoA kinase
M. conjunctivae (HAD-DPCK)	specific	3	264	1.60e-34	COG0561	Cof	cl26787	Hydroxymethylpyrimidine pyrophosphatase and other HAD family phosphatases [Coenzyme transport and metabolism, General function prediction only]
	superfamily	3	264	1.60e-34	cl26787	Hydrolase_3 superfamily	-	haloacid dehalogenase-like hydrolase
	non-specific	7	257	2.72e-33	pfam08282	Hydrolase_3	cl26787	haloacid dehalogenase-like hydrolase
	non-specific	6	259	1.10e-26	cd07516	HAD_Pase	cl21460	phosphatase, similar to Escherichia coli Cof and Thermotoga maritima TM0651
	superfamily	6	259	1.10e-26	cl21460	HAD_like superfamily	-	Haloacid Dehalogenase-like Hydrolases
	non-specific	6	257	2.46e-26	TIGR00099	Cof-subfamily	cl26787	Cof subfamily of IIB subfamily of haloacid dehalogenase superfamily
	non-specific	1	260	1.17e-13	PRK01158	PRK01158	cl26787	phosphoglycolate phosphatase
	non-specific	6	220	8.35e-13	TIGR01484	HAD-SF-IIB	cl26787	HAD-superfamily
	non-specific	10	263	2.26e-12	TIGR01482	SPP-subfamily	cl26787	sucrose-phosphate phosphatase subfamily
	non-specific	5	260	2.08e-11	cd07517	HAD_HPP	cl21460	phosphatase, similar to Bacteroides thetaiotaomicron VPI-5482 BT4131 hexose phosphate phosphatase
	non-specific	267	393	8.59e-10	COG0237	CoaE	cl28605	Dephospho-CoA kinase
	superfamily	267	393	8.59e-10	cl28605	CoaE superfamily	-	Dephospho-CoA kinase
	non-specific	267	401	2.23e-09	cd02022	DPCK	cl17190	Dephospho-CoA kinase
	superfamily	267	401	2.23e-09	cl17190	NK superfamily	-	Nucleoside/nucleotide kinase superfamily
	non-specific	5	257	5.30e-09	TIGR01487	Pglycolate_arch	cl26787	phosphoglycolate phosphatase, TA0175-type
	non-specific	5	261	9.68e-09	PRK10513	PRK10513	cl26787	sugar phosphate phosphatase
	non-specific	186	260	5.79e-08	cd07514	HAD_Pase	cl21460	phosphatase, similar to Thermoplasma acidophilum TA0175 phosphoglycolate phosphatase (PCPase), and Pyrococcus horikoshii PH1421, a magnesium-dependent phosphatase
	non-specific	267	414	1.65e-07	TIGR00152	TIGR00152	cl17190	Dephospho-CoA kinase
	non-specific	6	219	2.80e-07	PRK10976	PRK10976	cl26787	putative hydrolase
	non-specific	122	258	5.35e-05	cd07518	HAD_YbiV-Like	cl21460	Escherichia coli YbiV sugar phosphatase/phosphotransferase and related proteins
	non-specific	157	260	0.000276	cd02605	HAD_SPP	cl21460	sucrose-phosphatase, similar to Synechocystis sp PCC 6803 SPP
	non-specific	3	102	0.000756	COG1877	OtsB	cl28591	Trehalose-6-phosphatase [Carbohydrate transport and metabolism]
	superfamily	3	102	0.000756	cl28591	OtsB superfamily	-	Trehalose-6-phosphatase [Carbohydrate transport and metabolism]
	non-specific	1	76	0.001366	PRK00192	PRK00192	cl26786	mannosyl-3-phosphoglycerate phosphatase
	superfamily	1	76	0.001366	cl26786	YedP superfamily	-	Predicted mannosyl-3-phosphoglycerate phosphatase, HAD superfamily
M. cricetuli	non-specific	1	183	2.25e-17	COG0237	CoaE	cl28605	Dephospho-CoA kinase

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. cricetuli	superfamily	1	183	2.25e-17	cl28605	CoaE superfamily	-	Dephospho-CoA kinase
	non-specific	1	136	6.93e-13	cd02022	DPCK	cl17190	Dephospho-CoA kinase
	superfamily	1	136	6.93e-13	cl17190	NK superfamily	-	Nucleoside/nucleotide kinase superfamily
	non-specific	2	142	1.58e-08	TIGR00152	TIGR00152	cl17190	Dephospho-CoA kinase
	non-specific	2	136	3.11e-06	PRK00081	coaE	cl17190	Dephospho-CoA kinase
	non-specific	1	41	0.002388	COG1936	Fap7	cl17190	Broad-specificity NMP kinase
	non-specific	2	137	0.004333	COG0572	Udk	cl28606	Uridine kinase superfamily
	superfamily	2	137	0.004333	cl28606	Udk superfamily	-	Uridine kinase superfamily
M. crocodyli	non-specific	1	190	1.56e-15	COG0237	CoaE	cl28605	Dephospho-CoA kinase
	superfamily	1	190	1.56e-15	cl28605	CoaE superfamily	-	Dephospho-CoA kinase
	non-specific	1	137	1.08e-14	cd02022	DPCK	cl17190	Dephospho-CoA kinase
	superfamily	1	137	1.08e-14	cl17190	NK superfamily	-	Nucleoside/nucleotide kinase superfamily
	non-specific	2	173	6.02e-11	TIGR00152	TIGR00152	cl17190	Dephospho-CoA kinase
	non-specific	2	137	1.15e-08	PRK00081	coaE	cl17190	Dephospho-CoA kinase
	non-specific	2	137	3.04e-05	pfam01121	CoaE	cl17190	Dephospho-CoA kinase
M. dispar (HAD-DPCK)	non-specific	7	221	1.49e-29	pfam08282	Hydrolase_3	cl26787	haloacid dehalogenase-like hydrolase
	superfamily	7	221	1.49e-29	cl26787	Hydrolase_3 superfamily	-	haloacid dehalogenase-like hydrolase
	specific	1	266	4.21e-29	COG0561	Cof	cl26787	Hydroxymethylpyrimidine pyrophosphatase and other HAD family phosphatases [Coenzyme transport and metabolism, General function prediction only]
	non-specific	7	221	3.98e-24	cd07516	HAD_Pase	cl21460	phosphatase, similar to Escherichia coli Cof and Thermotoga maritima TM0651
	superfamily	7	221	3.98e-24	cl21460	HAD_like superfamily	-	Haloacid Dehalogenase-like Hydrolases
	non-specific	7	221	8.47e-23	TIGR00099	Cof-subfamily	cl26787	Cof subfamily of IIB subfamily of haloacid dehalogenase superfamily
	non-specific	7	221	5.10e-14	TIGR01484	HAD-SF-IIB	cl26787	HAD-superfamily
	non-specific	269	396	7.14e-13	COG0237	CoaE	cl28605	Dephospho-CoA kinase
	superfamily	269	396	7.14e-13	cl28605	CoaE superfamily	-	Dephospho-CoA kinase
	non-specific	4	108	1.70e-11	cd07517	HAD_HPP	cl21460	phosphatase, similar to Bacteroides thetaiotaomicron VPI-5482 BT4131 hexose phosphate phosphatase
	non-specific	269	396	3.50e-08	cd02022	DPCK	cl17190	Dephospho-CoA kinase
	superfamily	269	396	3.50e-08	cl17190	NK superfamily	-	Nucleoside/nucleotide kinase superfamily
	non-specific	1	221	4.95e-06	PRK01158	PRK01158	cl26787	phosphoglycolate phosphatase
	non-specific	269	428	8.47e-06	TIGR00152	TIGR00152	cl17190	Dephospho-CoA kinase
	non-specific	9	220	9.22e-06	TIGR01482	SPP-subfamily	cl26787	sucrose-phosphate phosphatase subfamily
	non-specific	9	220	8.38e-05	TIGR01487	Pglycolate_arch	cl26787	phosphoglycolate phosphatase, TA0175-type
	non-specific	2	92	0.000353	PRK10976	PRK10976	cl26787	putative hydrolase
	non-specific	7	49	0.000784	pfam02358	Trehalose_PPase	cl21460	Trehalose-phosphatase
	non-specific	7	48	0.001349	cd01627	HAD_TPP	cl21460	trehalose-phosphate phosphatase similar to Escherichia coli trehalose-6-phosphate phosphatase OtsB and Saccharomyces cerevisiae trehalose-phosphatase TPS2

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. dispar (HAD-DPCK)	non-specific	7	78	0.004988	COG1877	OtsB	cl28591	Trehalose-6-phosphatase [Carbohydrate transport and metabolism]
	superfamily	7	78	0.004988	cl28591	OtsB superfamily	-	Trehalose-6-phosphatase [Carbohydrate transport and metabolism]
M. felifaucium	non-specific	1	146	3.67e-16	COG0237	CoaE	cl28605	Dephospho-CoA kinase
	superfamily	1	146	3.67e-16	cl28605	CoaE superfamily	-	Dephospho-CoA kinase
	non-specific	1	136	4.80e-15	cd02022	DPCK	cl17190	Dephospho-CoA kinase
	superfamily	1	136	4.80e-15	cl17190	NK superfamily	-	Nucleoside/nucleotide kinase superfamily
	non-specific	2	186	8.51e-11	TIGR00152	TIGR00152	cl17190	Dephospho-CoA kinase
	non-specific	2	136	2.85e-07	PRK00081	coaE	cl17190	Dephospho-CoA kinase
	non-specific	2	137	0.003639	pfam01121	CoaE	cl17190	Dephospho-CoA kinase
M. felis	non-specific	3	122	3.08e-14	cd02022	DPCK	cl17190	Dephospho-CoA kinase
	superfamily	3	122	3.08e-14	cl17190	NK superfamily	-	Nucleoside/nucleotide kinase superfamily
	non-specific	1	107	2.19e-11	COG0237	CoaE	cl28605	Dephospho-CoA kinase
	superfamily	1	107	2.19e-11	cl28605	CoaE superfamily	-	Dephospho-CoA kinase
	non-specific	4	133	3.29e-08	TIGR00152	TIGR00152	cl17190	Dephospho-CoA kinase
	non-specific	1	122	6.46e-08	PRK00081	coaE	cl17190	Dephospho-CoA kinase
	non-specific	4	107	0.001021	pfam01121	CoaE	cl17190	Dephospho-CoA kinase
	non-specific	4	107	0.002914	PRK14732	coaE	cl17190	Dephospho-CoA kinase
	non-specific	1	26	0.003529	PRK15467	PRK15467	cl21455	ethanolamine utilization protein EutP
	superfamily	1	26	0.003529	cl21455	P-loop_NTPase superfamily	-	P-loop containing Nucleoside Triphosphate Hydrolases
	non-specific	2	40	0.00391	COG1936	Fap7	cl17190	Broad-specificity NMP kinase
M. fermentans	non-specific	1	146	7.45e-23	COG0237	CoaE	cl28605	Dephospho-CoA kinase
	superfamily	1	146	7.45e-23	cl28605	CoaE superfamily	-	Dephospho-CoA kinase
	specific	1	146	2.70e-20	cd02022	DPCK	cl17190	Dephospho-CoA kinase
	superfamily	1	146	2.70e-20	cl17190	NK superfamily	-	Nucleoside/nucleotide kinase superfamily
	non-specific	2	146	1.34e-14	PRK00081	coaE	cl17190	Dephospho-CoA kinase
	non-specific	2	146	8.01e-14	TIGR00152	TIGR00152	cl17190	Dephospho-CoA kinase
	non-specific	1	187	8.48e-07	PRK14732	coaE	cl17190	Dephospho-CoA kinase
	non-specific	2	137	3.08e-06	PRK14733	coaE	cl17190	Dephospho-CoA kinase
	non-specific	2	146	3.22e-06	pfam01121	CoaE	cl17190	Dephospho-CoA kinase
M. flocculare (HAD-DPCK)	non-specific	7	258	1.96e-37	pfam08282	Hydrolase_3	cl26787	haloacid dehalogenase-like hydrolase
	superfamily	7	258	1.96e-37	cl26787	Hydrolase_3 superfamily	-	haloacid dehalogenase-like hydrolase
	specific	1	263	3.22e-34	COG0561	Cof	cl26787	Hydroxymethylpyrimidine pyrophosphatase and other HAD family phosphatases [Coenzyme transport and metabolism, General function prediction only]
	non-specific	7	260	7.20e-31	cd07516	HAD_Pase	cl21460	phosphatase, similar to Escherichia coli Cof and Thermotoga maritima TM0651
	superfamily	7	260	7.20e-31	cl21460	HAD_like superfamily	-	Haloacid Dehalogenase-like Hydrolases
	non-specific	7	257	9.73e-30	TIGR00099	Cof-subfamily	cl26787	Cof subfamily of IIB subfamily of haloacid dehalogenase superfamily

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. flocculare (HAD-DPCK)	non-specific	4	246	5.58e-17	cd07517	HAD_HPP	cl21460	phosphatase, similar to Bacteroides thetaiotaomicron VPI-5482 BT4131 hexose phosphate phosphatase
	non-specific	268	395	1.17e-14	COG0237	CoaE	cl28605	Dephospho-CoA kinase
	superfamily	268	395	1.17e-14	cl28605	CoaE superfamily	-	Dephospho-CoA kinase
	non-specific	7	232	2.46e-12	TIGR01484	HAD-SF-IIB	cl26787	HAD-superfamily
	non-specific	1	246	5.34e-10	PRK01158	PRK01158	cl26787	phosphoglycolate phosphatase
	non-specific	268	395	1.46e-09	cd02022	DPCK	cl17190	Dephospho-CoA kinase
	superfamily	268	395	1.46e-09	cl17190	NK superfamily	-	Nucleoside/nucleotide kinase superfamily
	non-specific	9	250	4.25e-07	TIGR01482	SPP-subfamily	cl26787	sucrose-phosphate phosphatase subfamily
	non-specific	9	246	1.63e-06	TIGR01487	Pglycolate_arch	cl26787	phosphoglycolate phosphatase, TA0175-type
	non-specific	167	263	8.18e-05	PRK14502	PRK14502	cl26786	bifunctional mannosyl-3-phosphoglycerate synthase/mannosyl-3 phosphoglycerate phosphatase
	superfamily	167	263	8.18e-05	cl26786	YedP superfamily	-	Predicted mannosyl-3-phosphoglycerate phosphatase, HAD superfamily
	non-specific	1	263	0.00034	PRK10513	PRK10513	cl26787	sugar phosphate phosphatase
	non-specific	191	246	0.000356	cd07514	HAD_Pase	cl21460	phosphatase, similar to Thermoplasma acidophilum TA0175 phosphoglycolate phosphatase (PCPase), and Pyrococcus horikoshii PH1421, a magnesium-dependent phosphatase
	non-specific	268	405	0.000473	TIGR00152	TIGR00152	cl17190	Dephospho-CoA kinase
	non-specific	268	395	0.004105	PRK00081	coaE	cl17190	Dephospho-CoA kinase
	non-specific	246	285	0.004508	cd03257	ABC_NikE_OppD_transporters	cl25403	ATP-binding cassette domain of nickel/oligopeptides specific transporters
	superfamily	246	285	0.004508	cl25403	ABC_ATPase superfamily	-	ATP-binding cassette transporter nucleotide-binding domain
	non-specific	7	74	0.006724	cd01627	HAD_TPP	cl21460	trehalose-phosphate phosphatase similar to Escherichia coli trehalose-6-phosphate phosphatase OtsB and Saccharomyces cerevisiae trehalose-phosphatase TPS2
M. gallinarum	non-specific	1	145	3.41e-15	COG0237	CoaE	cl28605	Dephospho-CoA kinase
	superfamily	1	145	3.41e-15	cl28605	CoaE superfamily	-	Dephospho-CoA kinase
	non-specific	1	136	4.53e-13	cd02022	DPCK	cl17190	Dephospho-CoA kinase
	superfamily	1	136	4.53e-13	cl17190	NK superfamily	-	Nucleoside/nucleotide kinase superfamily
	non-specific	2	145	2.67e-09	TIGR00152	TIGR00152	cl17190	Dephospho-CoA kinase
	non-specific	2	151	2.01e-05	PRK00081	coaE	cl17190	Dephospho-CoA kinase
M. gallisepticum	non-specific	25	199	6.87e-19	COG0237	CoaE	cl28605	Dephospho-CoA kinase
	superfamily	25	199	6.87e-19	cl28605	CoaE superfamily	-	Dephospho-CoA kinase
	non-specific	24	194	4.11e-18	PRK00081	coaE	cl17190	Dephospho-CoA kinase
	superfamily	24	194	4.11e-18	cl17190	NK superfamily	-	Nucleoside/nucleotide kinase superfamily
	non-specific	23	186	9.11e-17	TIGR00152	TIGR00152	cl17190	Dephospho-CoA kinase
	non-specific	17	181	6.24e-12	cd02022	DPCK	cl17190	Dephospho-CoA kinase

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. gallisepticum	non-specific	23	88	5.55e-08	pfam01121	CoaE	cl17190	Dephospho-CoA kinase
	non-specific	18	195	2.61e-07	PLN02422	PLN02422	cl17190	Dephospho-CoA kinase
	non-specific	27	194	3.52e-07	PRK14732	coaE	cl17190	Dephospho-CoA kinase
M. genitalium	specific	2	182	5.10e-78	TIGR00152	TIGR00152	cl17190	Dephospho-CoA kinase
	superfamily	2	182	5.10e-78	cl17190	NK superfamily	-	Nucleoside/nucleotide kinase superfamily
	specific	1	198	1.74e-45	COG0237	CoaE	cl28605	Dephospho-CoA kinase
	superfamily	1	198	1.74e-45	cl28605	CoaE superfamily	-	Dephospho-CoA kinase
	specific	2	169	1.84e-37	cd02022	DPCK	cl17190	Dephospho-CoA kinase
	non-specific	3	128	1.43e-18	PRK00081	coaE	cl17190	Dephospho-CoA kinase
	non-specific	3	106	8.84e-10	PRK14730	coaE	cl17190	Dephospho-CoA kinase
	non-specific	1	110	8.42e-08	pfam01121	CoaE	cl17190	Dephospho-CoA kinase
	non-specific	2	98	2.74e-05	PLN02422	PLN02422	cl17190	Dephospho-CoA kinase
	non-specific	2	114	9.20e-05	PRK14732	coaE	cl17190	Dephospho-CoA kinase
	non-specific	3	75	0.000209	pfam13521	AAA_28	cl21455	AAA domain
	superfamily	3	75	0.000209	cl21455	P-loop_NTPase superfamily	-	P-loop containing Nucleoside Triphosphate Hydrolases
	non-specific	1	172	0.000756	PTZ00451	PTZ00451	cl17190	Dephospho-CoA kinase
	non-specific	6	77	0.002063	pfam13207	AAA_17	cl21455	AAA domain
	non-specific	2	82	0.002657	PRK14731	coaE	cl17190	Dephospho-CoA kinase
	non-specific	3	50	0.003221	TIGR00231	small_GTP	cl27769	small GTP-binding protein domain
	superfamily	3	50	0.003221	cl27769	GTP_EFTU superfamily	-	Elongation factor Tu GTP binding domain
M. hyopneumoniae (HAD-DPCK)	specific	1	265	8.37e-32	COG0561	Cof	cl26787	Hydroxymethylpyrimidine pyrophosphatase and other HAD family phosphatases [Coenzyme transport and metabolism, General function prediction only]
	superfamily	1	265	8.37e-32	cl26787	Hydrolase_3 superfamily	-	haloacid dehalogenase-like hydrolase
	non-specific	7	258	1.25e-29	pfam08282	Hydrolase_3	cl26787	haloacid dehalogenase-like hydrolase
	non-specific	5	258	3.30e-27	TIGR00099	Cof-subfamily	cl26787	Cof subfamily of IIB subfamily of haloacid dehalogenase superfamily
	non-specific	7	260	9.11e-25	cd07516	HAD_Pase	cl21460	phosphatase, similar to Escherichia coli Cof and Thermotoga maritima TM0651
	superfamily	7	260	9.11e-25	cl21460	HAD_like superfamily	-	Haloacid Dehalogenase-like hydrolases
	non-specific	269	428	1.87e-13	COG0237	CoaE	cl28605	Dephospho-CoA kinase
	superfamily	269	428	1.87e-13	cl28605	CoaE superfamily	-	Dephospho-CoA kinase
	non-specific	4	246	2.03e-12	cd07517	HAD_HPP	cl21460	phosphatase, similar to Bacteroides thetaiotaomicron VPI-5482 BT4131 hexose phosphate phosphatase
	non-specific	7	221	3.80e-11	TIGR01484	HAD-SF-IIB	cl26787	HAD-superfamily
	non-specific	9	250	3.14e-09	TIGR01482	SPP-subfamily	cl26787	sucrose-phosphate phosphatase subfamily
	non-specific	269	395	1.51e-08	cd02022	DPCK	cl17190	Dephospho-CoA kinase
	superfamily	269	395	1.51e-08	cl17190	NK superfamily	-	Nucleoside/nucleotide kinase superfamily
	non-specific	1	263	3.41e-08	PRK01158	PRK01158	cl26787	phosphoglycolate phosphatase
	non-specific	269	428	2.04e-07	TIGR00152	TIGR00152	cl17190	Dephospho-CoA kinase

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. hyopneumoniae (HAD-DPCK)	non-specific	191	246	1.35e-06	cd07514	HAD_Pase	cl21460	phosphatase, similar to Thermoplasma acidophilum TA0175 phosphoglycolate phosphatase (PCPase), and Pyrococcus horikoshii PH1421, a magnesium-dependent phosphatase
	non-specific	7	215	6.16e-06	pfam02358	Trehalose_PPase	cl21460	Trehalose-phosphatase
	non-specific	9	246	3.75e-05	TIGR01487	Pglycolate_arch	cl26787	phosphoglycolate phosphatase, TA0175-type
	non-specific	7	48	0.000556	cd01627	HAD_TPP	cl21460	trehalose-phosphate phosphatase similar to Escherichia coli trehalose-6-phosphate phosphatase OtsB and Saccharomyces cerevisiae trehalose-phosphatase TPS2
	non-specific	182	258	0.000585	PLN02887	PLN02887	cl26787	hydrolase family protein
	non-specific	183	263	0.001299	PRK10513	PRK10513	cl26787	sugar phosphate phosphatase
	non-specific	7	58	0.002792	TIGR00685	T6PP	cl21460	trehalose-phosphatase
	non-specific	183	258	0.009025	cd07518	HAD_YbiV-Like	cl21460	Escherichia coli YbiV sugar phosphatase/phosphotransferase and related proteins
M. hyorhinis	non-specific	9	141	1.14e-15	COG0237	CoaE	cl28605	Dephospho-CoA kinase
	superfamily	9	141	1.14e-15	cl28605	CoaE superfamily	-	Dephospho-CoA kinase
	non-specific	12	141	2.74e-13	cd02022	DPCK	cl17190	Dephospho-CoA kinase
	superfamily	12	141	2.74e-13	cl17190	NK superfamily	-	Nucleoside/nucleotide kinase superfamily
	non-specific	9	140	1.34e-08	TIGR00152	TIGR00152	cl17190	Dephospho-CoA kinase
	non-specific	12	141	5.89e-06	PRK00081	coaE	cl17190	Dephospho-CoA kinase
M. imitans	specific	7	196	6.69e-24	COG0237	CoaE	cl28605	Dephospho-CoA kinase
	superfamily	7	196	6.69e-24	cl28605	CoaE superfamily	-	Dephospho-CoA kinase
	non-specific	9	193	6.46e-23	TIGR00152	TIGR00152	cl17190	Dephospho-CoA kinase
	superfamily	9	193	6.46e-23	cl17190	NK superfamily	-	Nucleoside/nucleotide kinase superfamily
	specific	9	182	3.99e-20	cd02022	DPCK	cl17190	Dephospho-CoA kinase
	non-specific	12	194	1.25e-19	PRK00081	coaE	cl17190	Dephospho-CoA kinase
	non-specific	9	196	1.11e-11	PLN02422	PLN02422	cl17190	Dephospho-CoA kinase
	non-specific	8	90	3.95e-11	pfam01121	CoaE	cl17190	Dephospho-CoA kinase
	non-specific	9	97	2.22e-05	PRK14732	coaE	cl17190	Dephospho-CoA kinase
	non-specific	12	194	2.64e-05	PRK14730	coaE	cl17190	Dephospho-CoA kinase
M. iners	non-specific	1	187	2.66e-13	COG0237	CoaE	cl28605	Dephospho-CoA kinase
	superfamily	1	187	2.66e-13	cl28605	CoaE superfamily	-	Dephospho-CoA kinase
	non-specific	1	160	5.90e-12	cd02022	DPCK	cl17190	Dephospho-CoA kinase
	superfamily	1	160	5.90e-12	cl17190	NK superfamily	-	Nucleoside/nucleotide kinase superfamily
	non-specific	2	100	5.22e-07	TIGR00152	TIGR00152	cl17190	Dephospho-CoA kinase
	non-specific	1	20	4.04e-05	cd03213	ABCG_EPDR	cl25403	Eye pigment and drug resistance transporter subfamily G of the ATP-binding cassette superfamily
	superfamily	1	20	4.04e-05	cl25403	ABC_ATPase superfamily	-	ATP-binding cassette transporter nucleotide-binding domain
	non-specific	2	123	6.79e-05	PRK00081	coaE	cl17190	Dephospho-CoA kinase

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. iners	non-specific	1	19	0.000177	cd03255	ABC_MJ0796_LolCDE_FtsE	cl25403	ATP-binding cassette domain of the transporters involved in export of lipoprotein and macrolide, and cell division protein
	non-specific	1	22	0.000252	pfam00005	ABC_tran	cl21455	ABC transporter
	superfamily	1	22	0.000252	cl21455	P-loop_NTPase superfamily	-	P-loop containing Nucleoside Triphosphate Hydrolases
	non-specific	2	22	0.000398	pfam01926	MMR_HSR1	cl21455	50S ribosome-binding GTPase
	non-specific	1	17	0.000429	COG1136	LolD	cl28181	ABC-type lipoprotein export system, ATPase component [Cell wall/membrane/envelope biogenesis]
	superfamily	1	17	0.000429	cl28181	AAA superfamily	-	ATPases associated with a variety of cellular activities
	non-specific	1	16	0.000478	cd03256	ABC_PhnC_transporter	cl25403	ATP-binding cassette domain of the binding protein-dependent phosphonate transport system
	non-specific	1	20	0.000505	PRK10771	thiQ	cl28181	thiamine transporter ATP-binding subunit
	non-specific	1	21	0.000512	cd03228	ABCC_MRP_Like	cl25403	ATP-binding cassette domain of multidrug resistance protein-like transporters
	non-specific	2	23	0.000831	PRK11174	PRK11174	cl26602	cysteine/glutathione ABC transporter membrane/ATP-binding component
	superfamily	2	23	0.000831	cl26602	SunT superfamily	-	ABC-type bacteriocin/lantibiotic exporters, contain an N-terminal double-glycine peptidase domain [Defense mechanisms]
	non-specific	2	28	0.001139	COG4987	CydC	cl26602	ABC-type transport system involved in cytochrome bd biosynthesis, fused ATPase and permease components [Energy production and conversion, Posttranslational modification, protein turnover, chaperones]
	non-specific	2	20	0.001175	COG1116	TauB	cl28181	ABC-type nitrate/sulfonate/bicarbonate transport system, ATPase component [Inorganic ion transport and metabolism]
	non-specific	1	19	0.001453	TIGR02857	CydD	cl26602	thiol reductant ABC exporter, CydD subunit
	non-specific	2	19	0.001626	COG4988	CydD	cl26602	ABC-type transport system involved in cytochrome bd biosynthesis, ATPase and permease components [Energy production and conversion, Posttranslational modification, protein turnover, chaperones]
	non-specific	2	133	0.001635	pfam01121	CoaE	cl17190	Dephospho-CoA kinase
	non-specific	1	48	0.001853	COG3638	PhnC	cl28181	ABC-type phosphate/phosphonate transport system, ATPase component [Inorganic ion transport and metabolism]
	non-specific	1	19	0.002362	PRK10535	PRK10535	cl28180	macrolide transporter ATP-binding /permease protein

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. iners	superfamily	1	19	0.002362	cl28180	PRK10535 superfamily	-	macrolide transporter ATP-binding /permease protein
	non-specific	1	21	0.002445	COG1135	AbcC	cl28181	ABC-type methionine transport system, ATPase component [Amino acid transport and metabolism]
	non-specific	1	22	0.002501	COG0488	Uup	cl28181	ATPase components of ABC transporters with duplicated ATPase domains [General function prediction only]
	non-specific	3	22	0.002813	cd00880	Era_like	cl21455	E. coli Ras-like protein (Era)-like GTPase
	non-specific	2	28	0.003096	cd03247	ABCC_cytochrome_bd	cl25403	ATP-binding cassette domain of CydCD, subfamily C
	non-specific	3	52	0.00316	COG1119	ModF	cl28181	ABC-type molybdenum transport system, ATPase component/photorepair protein PhrA [Inorganic ion transport and metabolism]
	non-specific	2	61	0.003567	pfam01712	dNK	cl17190	Deoxynucleoside kinase
	non-specific	1	20	0.00391	TIGR00955	3a01204	cl28180	The Eye Pigment Precursor Transporter (EPP) Family protein
	non-specific	2	23	0.004193	COG3840	ThiQ	cl28181	ABC-type thiamine transport system, ATPase component [Coenzyme transport and metabolism]
	non-specific	1	19	0.004303	TIGR03608	L_ocin_972_ABC	cl28181	putative bacteriocin export ABC transporter, lactococcin 972 group
	non-specific	2	19	0.004448	PRK01889	PRK01889	cl26332	GTPase RsgA
	superfamily	2	19	0.004448	cl26332	DUF258 superfamily	-	Protein of unknown function, DUF258
	non-specific	1	20	0.004705	COG2274	SunT	cl26602	ABC-type bacteriocin/lantibiotic exporters, contain an N-terminal double-glycine peptidase domain [Defense mechanisms]
	non-specific	1	19	0.005264	TIGR02211	LolD_lipo_ex	cl28181	lipoprotein releasing system, ATP-binding protein
	non-specific	2	22	0.005457	COG0486	MnmE	cl26334	tRNA U34 5-carboxymethylaminomethyl modifying GTPase MnmE/TrmE [Translation, ribosomal structure and biogenesis]
	superfamily	2	22	0.005457	cl26334	MnmE_helical superfamily	-	MnmE helical domain
	non-specific	2	29	0.005518	PRK07261	PRK07261	cl25401	topology modulation protein
	superfamily	2	29	0.005518	cl25401	DEXDc superfamily	-	DEAD-like helicases superfamily
	non-specific	2	22	0.005561	cd04164	trmE	cl21455	trmE is a tRNA modification GTPase
	non-specific	2	20	0.005687	cd03293	ABC_NrtD_SsuB_transporters	cl25403	ATP-binding cassette domain of the nitrate and sulfonate transporters
	non-specific	1	17	0.006201	cd00267	ABC_ATPase	cl25403	ATP-binding cassette transporter nucleotide-binding domain
	non-specific	4	21	0.006699	cd03258	ABC_MetN_methionine_transporter	cl25403	ATP-binding cassette domain of methionine transporter
	non-specific	1	20	0.006877	PRK13548	hmuV	cl28181	hemin importer ATP-binding subunit

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. iners	non-specific	2	19	0.006969	cd01854	YjeQ_EngC	cl25406	Ribosomal interacting GTPase YjeQ/EngC, a circularly permuted subfamily of the Ras GTPases
	superfamily	2	19	0.006969	cl25406	YlqF_related_GTPase superfamily	-	Circularly permuted YlqF-related GTPases
	non-specific	1	17	0.00716	cd03225	ABC_cobalt_CbiO_domain1	cl25403	First domain of the ATP-binding cassette component of cobalt transport system
	non-specific	3	19	0.007294	TIGR00929	VirB4_CagE	cl26286	type IV secretion/conjugal transfer ATPase, VirB4 family
	superfamily	3	19	0.007294	cl26286	VirB4 superfamily	-	Type IV secretory pathway, VirB4 component [Intracellular trafficking, secretion, and vesicular transport]
	non-specific	1	19	0.007436	cd03229	ABC_Class3	cl25403	ATP-binding cassette domain of the binding protein-dependent transport systems
	non-specific	2	20	0.00811	TIGR02868	CydC	cl26602	thiol reductant ABC exporter, CydC subunit
	non-specific	2	22	0.008138	PRK05291	trmE	cl26334	tRNA modification GTPase TrmE
	non-specific	1	21	0.009792	cd03251	ABCC_MsbA	cl25403	ATP-binding cassette domain of the bacterial lipid flippase and related proteins, subfamily C
M. iowae	specific	10	199	4.84e-26	COG0237	CoaE	cl28605	Dephospho-CoA kinase
	superfamily	10	199	4.84e-26	cl28605	CoaE superfamily	-	Dephospho-CoA kinase
	specific	11	189	6.65e-25	cd02022	DPCK	cl17190	Dephospho-CoA kinase
	superfamily	11	189	6.65e-25	cl17190	NK superfamily	-	Nucleoside/nucleotide kinase superfamily
	non-specific	21	201	4.13e-20	PRK00081	coaE	cl17190	Dephospho-CoA kinase
	non-specific	11	199	3.87e-18	TIGR00152	TIGR00152	cl17190	Dephospho-CoA kinase
	non-specific	11	189	1.37e-11	pfam01121	CoaE	cl17190	Dephospho-CoA kinase
	non-specific	10	199	1.22e-08	PLN02422	PLN02422	cl17190	Dephospho-CoA kinase
	non-specific	10	199	5.33e-06	PRK14730	coaE	cl17190	Dephospho-CoA kinase
	non-specific	11	203	7.07e-06	PRK14732	coaE	cl17190	Dephospho-CoA kinase
	non-specific	10	199	0.000192	PRK01184	PRK01184	cl17190	hypothetical protein
	non-specific	10	201	0.002715	PRK14734	coaE	cl17190	Dephospho-CoA kinase
M. leachii	specific	2	175	4.88e-34	cd02022	DPCK	cl17190	Dephospho-CoA kinase
	superfamily	2	175	4.88e-34	cl17190	NK superfamily	-	Nucleoside/nucleotide kinase superfamily
	non-specific	1	168	3.69e-21	PRK00081	coaE	cl17190	Dephospho-CoA kinase
	non-specific	1	155	8.39e-19	COG0237	CoaE	cl28605	Dephospho-CoA kinase
	superfamily	1	155	8.39e-19	cl28605	CoaE superfamily	-	Dephospho-CoA kinase
	non-specific	1	163	1.02e-16	pfam01121	CoaE	cl17190	Dephospho-CoA kinase
	non-specific	2	170	1.28e-13	TIGR00152	TIGR00152	cl17190	Dephospho-CoA kinase
	non-specific	2	163	1.57e-10	PRK14732	coaE	cl17190	Dephospho-CoA kinase
	non-specific	2	185	2.36e-10	PRK14730	coaE	cl17190	Dephospho-CoA kinase
	non-specific	3	144	1.59e-08	PRK14733	coaE	cl17190	Dephospho-CoA kinase
	non-specific	2	161	0.000358	PRK14731	coaE	cl17190	Dephospho-CoA kinase
	non-specific	2	168	0.000801	PLN02422	PLN02422	cl17190	Dephospho-CoA kinase
M. leonicaptivi	non-specific	1	119	5.92e-10	cd02022	DPCK	cl17190	Dephospho-CoA kinase
	superfamily	1	119	5.92e-10	cl17190	NK superfamily	-	Nucleoside/nucleotide kinase superfamily

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. leonicaptivi	non-specific	1	37	2.16e-08	COG0237	CoaE	cl28605	Dephospho-CoA kinase
	superfamily	1	37	2.16e-08	cl28605	CoaE superfamily	-	Dephospho-CoA kinase
	non-specific	2	106	4.00e-07	PRK00081	coaE	cl17190	Dephospho-CoA kinase
	non-specific	2	174	0.000202	TIGR00152	TIGR00152	cl17190	Dephospho-CoA kinase
	non-specific	2	119	0.00902	pfam01121	CoaE	cl17190	Dephospho-CoA kinase
M. lipofaciens	non-specific	1	187	2.08e-18	COG0237	CoaE	cl28605	Dephospho-CoA kinase
	superfamily	1	187	2.08e-18	cl28605	CoaE superfamily	-	Dephospho-CoA kinase
	non-specific	1	137	3.80e-13	cd02022	DPCK	cl17190	Dephospho-CoA kinase
	superfamily	1	137	3.80e-13	cl17190	NK superfamily	-	Nucleoside/nucleotide kinase superfamily
	non-specific	2	137	7.85e-12	TIGR00152	TIGR00152	cl17190	Dephospho-CoA kinase
	non-specific	11	137	2.89e-06	PRK00081	coaE	cl17190	Dephospho-CoA kinase
M. mobile	non-specific	1	147	2.34e-16	COG0237	CoaE	cl28605	Dephospho-CoA kinase
	superfamily	1	147	2.34e-16	cl28605	CoaE superfamily	-	Dephospho-CoA kinase
	non-specific	1	147	1.90e-15	cd02022	DPCK	cl17190	Dephospho-CoA kinase
	superfamily	1	147	1.90e-15	cl17190	NK superfamily	-	Nucleoside/nucleotide kinase superfamily
	non-specific	2	172	3.00e-10	TIGR00152	TIGR00152	cl17190	Dephospho-CoA kinase
	non-specific	2	147	9.86e-09	PRK00081	coaE	cl17190	Dephospho-CoA kinase
	non-specific	1	32	0.000576	COG1936	Fap7	cl17190	Broad-specificity NMP kinase
	non-specific	2	147	0.006349	pfam01121	CoaE	cl17190	Dephospho-CoA kinase
M. molare	non-specific	1	136	4.12e-19	COG0237	CoaE	cl28605	Dephospho-CoA kinase
	superfamily	1	136	4.12e-19	cl28605	CoaE superfamily	-	Dephospho-CoA kinase
	non-specific	1	131	3.33e-14	cd02022	DPCK	cl17190	Dephospho-CoA kinase
	superfamily	1	131	3.33e-14	cl17190	NK superfamily	-	Nucleoside/nucleotide kinase superfamily
	non-specific	2	131	1.64e-09	TIGR00152	TIGR00152	cl17190	Dephospho-CoA kinase
	non-specific	2	131	1.16e-05	PRK00081	coaE	cl17190	Dephospho-CoA kinase
M. mycoides subsp. *capri*	specific	2	175	2.52e-36	cd02022	DPCK	cl17190	Dephospho-CoA kinase
	superfamily	2	175	2.52e-36	cl17190	NK superfamily	-	Nucleoside/nucleotide kinase superfamily
	non-specific	1	170	4.86e-24	PRK00081	coaE	cl17190	Dephospho-CoA kinase
	non-specific	1	175	9.60e-21	COG0237	CoaE	cl28605	Dephospho-CoA kinase
	superfamily	1	175	9.60e-21	cl28605	CoaE superfamily	-	Dephospho-CoA kinase
	non-specific	1	170	9.20e-20	pfam01121	CoaE	cl17190	Dephospho-CoA kinase
	non-specific	2	170	5.02e-15	TIGR00152	TIGR00152	cl17190	Dephospho-CoA kinase
	non-specific	2	170	3.85e-12	PRK14730	coaE	cl17190	Dephospho-CoA kinase
	non-specific	2	170	5.65e-12	PRK14732	coaE	cl17190	Dephospho-CoA kinase
	non-specific	3	170	6.95e-09	PRK14733	coaE	cl17190	Dephospho-CoA kinase
	non-specific	2	170	1.76e-06	PRK14731	coaE	cl17190	Dephospho-CoA kinase
	non-specific	2	175	3.12e-05	PLN02422	PLN02422	cl17190	Dephospho-CoA kinase
M. mycoides subsp. *mycoides*	specific	6	179	5.87e-34	cd02022	DPCK	cl17190	Dephospho-CoA kinase
	superfamily	6	179	5.87e-34	cl17190	NK superfamily	-	Nucleoside/nucleotide kinase superfamily
	non-specific	4	172	6.04e-21	PRK00081	coaE	cl17190	Dephospho-CoA kinase
	non-specific	4	159	7.48e-19	COG0237	CoaE	cl28605	Dephospho-CoA kinase
	superfamily	4	159	7.48e-19	cl28605	CoaE superfamily	-	Dephospho-CoA kinase
	non-specific	5	167	9.70e-17	pfam01121	CoaE	cl17190	Dephospho-CoA kinase

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. mycoides subsp. *mycoides*	non-specific	6	174	9.44e-14	TIGR00152	TIGR00152	cl17190	Dephospho-CoA kinase
	non-specific	6	167	9.65e-11	PRK14732	coaE	cl17190	Dephospho-CoA kinase
	non-specific	6	189	9.75e-11	PRK14730	coaE	cl17190	Dephospho-CoA kinase
	non-specific	7	148	3.64e-08	PRK14733	coaE	cl17190	Dephospho-CoA kinase
	non-specific	3	165	0.000422	PRK14731	coaE	cl17190	Dephospho-CoA kinase
	non-specific	6	172	0.001703	PLN02422	PLN02422	cl17190	Dephospho-CoA kinase
M. opalescens	non-specific	1	135	8.34e-17	COG0237	CoaE	cl28605	Dephospho-CoA kinase
	superfamily	1	135	8.34e-17	cl28605	CoaE superfamily	-	Dephospho-CoA kinase
	non-specific	1	135	3.58e-11	cd02022	DPCK	cl17190	Dephospho-CoA kinase
	superfamily	1	135	3.58e-11	cl17190	NK superfamily	-	Nucleoside/nucleotide kinase superfamily
	non-specific	2	111	7.67e-09	PRK00081	coaE	cl17190	Dephospho-CoA kinase
	non-specific	2	159	3.30e-08	TIGR00152	TIGR00152	cl17190	Dephospho-CoA kinase
	non-specific	2	129	1.34e-06	pfam01121	CoaE	cl17190	Dephospho-CoA kinase
	non-specific	2	33	0.002374	cd01983	Fer4_NifH	cl28886	The Fer4_NifH superfamily
	superfamily	2	33	0.002374	cl28886	Fer4_NifH superfamily	-	The Fer4_NifH superfamily
	non-specific	2	21	0.003859	COG4598	HisP	cl28181	ABC-type histidine transport system, ATPase component [Amino acid transport and metabolism]
	superfamily	2	21	0.003859	cl28181	AAA superfamily	-	ATPases associated with a variety of cellular activities
	non-specific	2	135	0.007313	PRK14730	coaE	cl17190	Dephospho-CoA kinase
M. ovipneumoniae (HAD-DPCK)	non-specific	5	259	3.75e-35	TIGR00099	Cof-subfamily	cl26787	Cof subfamily of IIB subfamily of haloacid dehalogenase superfamily
	superfamily	5	259	3.75e-35	cl26787	Hydrolase_3 superfamily	-	haloacid dehalogenase-like hydrolase
	specific	1	266	3.19e-34	COG0561	Cof	cl26787	Hydroxymethylpyrimidine pyrophosphatase and other HAD family phosphatases [Coenzyme transport and metabolism, General function prediction only]
	non-specific	7	259	4.26e-34	pfam08282	Hydrolase_3	cl26787	haloacid dehalogenase-like hydrolase
	non-specific	5	261	1.17e-27	cd07516	HAD_Pase	cl21460	phosphatase, similar to Escherichia coli Cof and Thermotoga maritima TM0651
	superfamily	5	261	1.17e-27	cl21460	HAD_like superfamily	-	Haloacid Dehalogenase-like Hydrolases
	non-specific	4	262	2.92e-18	cd07517	HAD_HPP	cl21460	phosphatase, similar to Bacteroides thetaiotaomicron VPI-5482 BT4131 hexose phosphate phosphatase
	non-specific	269	394	4.40e-12	COG0237	CoaE	cl28605	Dephospho-CoA kinase
	superfamily	269	394	4.40e-12	cl28605	CoaE superfamily	-	Dephospho-CoA kinase
	non-specific	269	394	8.41e-09	cd02022	DPCK	cl17190	Dephospho-CoA kinase
	superfamily	269	394	8.41e-09	cl17190	NK superfamily	-	Nucleoside/nucleotide kinase superfamily
	non-specific	5	233	9.07e-09	TIGR01484	HAD-SF-IIB	cl26787	HAD-superfamily
	non-specific	1	262	1.87e-08	PRK01158	PRK01158	cl26787	phosphoglycolate phosphatase
	non-specific	8	251	5.77e-08	TIGR01482	SPP-subfamily	cl26787	sucrose-phosphate phosphatase subfamily
	non-specific	269	366	1.92e-07	TIGR00152	TIGR00152	cl17190	Dephospho-CoA kinase
	non-specific	1	265	1.37e-06	PRK10513	PRK10513	cl26787	sugar phosphate phosphatase

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. ovipneumoniae (HAD-DPCK)	non-specific	4	259	0.001224	TIGR01487	Pglycolate_arch	cl26787	phosphoglycolate phosphatase, TA0175-type
	non-specific	263	438	0.004457	PRK14733	coaE	cl17190	Dephospho-CoA kinase
	non-specific	7	74	0.007687	pfam02358	Trehalose_PPase	cl21460	Trehalose-phosphatase
M. penetrans	specific	14	143	1.58e-39	cd02022	DPCK	cl17190	Dephospho-CoA kinase
	superfamily	14	143	1.58e-39	cl17190	NK superfamily	-	Nucleoside/nucleotide kinase superfamily
	specific	13	142	1.53e-25	COG0237	CoaE	cl28605	Dephospho-CoA kinase
	superfamily	13	142	1.53e-25	cl28605	CoaE superfamily	-	Dephospho-CoA kinase
	non-specific	12	142	3.04e-24	PRK00081	coaE	cl17190	Dephospho-CoA kinase
	non-specific	14	143	2.66e-19	TIGR00152	TIGR00152	cl17190	Dephospho-CoA kinase
	non-specific	14	140	8.47e-15	pfam01121	CoaE	cl17190	Dephospho-CoA kinase
	non-specific	14	203	2.35e-11	PRK14731	coaE	cl17190	Dephospho-CoA kinase
	non-specific	15	142	1.05e-09	PRK14730	coaE	cl17190	Dephospho-CoA kinase
	non-specific	13	106	4.78e-09	PLN02422	PLN02422	cl17190	Dephospho-CoA kinase
	non-specific	14	203	1.78e-08	PRK14732	coaE	cl17190	Dephospho-CoA kinase
	non-specific	11	141	5.77e-05	PRK14733	coaE	cl17190	Dephospho-CoA kinase
	non-specific	13	143	0.000212	PRK14734	coaE	cl17190	Dephospho-CoA kinase
	non-specific	15	143	0.000749	pfam13238	AAA_18	cl21455	AAA domain
	superfamily	15	143	0.000749	cl21455	P-loop_NTPase superfamily	-	P-loop containing Nucleoside Triphosphate Hydrolases
	non-specific	13	30	0.001651	PRK00349	uvrA	cl26603	excinuclease ABC subunit A
	superfamily	13	30	0.001651	cl26603	UvrA superfamily	-	Excinuclease UvrABC ATPase subunit [Replication, recombination and repair]
M. pirum	specific	4	194	9.22e-25	COG0237	CoaE	cl28605	Dephospho-CoA kinase
	superfamily	4	194	9.22e-25	cl28605	CoaE superfamily	-	Dephospho-CoA kinase
	non-specific	7	187	1.24e-22	TIGR00152	TIGR00152	cl17190	Dephospho-CoA kinase
	superfamily	7	187	1.24e-22	cl17190	NK superfamily	-	Nucleoside/nucleotide kinase superfamily
	specific	7	179	2.50e-22	cd02022	DPCK	cl17190	Dephospho-CoA kinase
	non-specific	6	192	1.75e-15	PRK00081	coaE	cl17190	Dephospho-CoA kinase
	non-specific	6	92	3.16e-09	pfam01121	CoaE	cl17190	Dephospho-CoA kinase
	non-specific	7	92	6.17e-07	PRK14730	coaE	cl17190	Dephospho-CoA kinase
	non-specific	7	92	0.000103	PRK14732	coaE	cl17190	Dephospho-CoA kinase
	non-specific	7	92	0.0002	COG1136	LolD	cl28181	ABC-type lipoprotein export system, ATPase component [Cell wall/membrane/envelope biogenesis]
	superfamily	7	92	0.0002	cl28181	AAA superfamily	-	ATPases associated with a variety of cellular activities
	non-specific	7	85	0.000317	PLN02422	PLN02422	cl17190	Dephospho-CoA kinase
	non-specific	7	25	0.000701	cd01854	YjeQ_EngC	cl25406	Ribosomal interacting GTPase YjeQ/EngC, a circularly permuted subfamily of the Ras GTPases
	superfamily	7	25	0.000701	cl25406	YlqF_related_GTPase superfamily	-	Circularly permuted YlqF-related GTPases
	non-specific	8	137	0.000793	pfam13238	AAA_18	cl21455	AAA domain

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. pirum	superfamily	8	137	0.000793	cl21455	P-loop_NTPase superfamily	-	P-loop containing Nucleoside Triphosphate Hydrolases
	non-specific	1	91	0.00103	PRK14731	coaE	cl17190	Dephospho-CoA kinase
	non-specific	6	27	0.001255	COG1936	Fap7	cl17190	Broad-specificity NMP kinase
	non-specific	6	33	0.002034	PRK04040	PRK04040	cl17190	adenylate kinase
	non-specific	7	92	0.002161	cd03255	ABC_MJ0796_LolCDE_FtsE	cl25403	ATP-binding cassette domain of the transporters involved in export of lipoprotein and macrolide, and cell division protein
	superfamily	7	92	0.002161	cl25403	ABC_ATPase superfamily	-	ATP-binding cassette transporter nucleotide-binding domain
	non-specific	1	28	0.003735	COG2019	AdkA	cl17190	Archaeal adenylate kinase
	non-specific	6	35	0.006283	PRK03839	PRK03839	cl17190	putative kinase
M. pneumoniae	specific	2	183	1.28e-81	TIGR00152	TIGR00152	cl17190	Dephospho-CoA kinase
	superfamily	2	183	1.28e-81	cl17190	NK superfamily	-	Nucleoside/nucleotide kinase superfamily
	specific	1	194	4.42e-48	COG0237	CoaE	cl28605	Dephospho-CoA kinase
	superfamily	1	194	4.42e-48	cl28605	CoaE superfamily	-	Dephospho-CoA kinase
	specific	2	156	1.15e-39	cd02022	DPCK	cl17190	Dephospho-CoA kinase
	non-specific	3	156	1.67e-28	PRK00081	coaE	cl17190	Dephospho-CoA kinase
	non-specific	3	106	3.23e-14	PRK14730	coaE	cl17190	Dephospho-CoA kinase
	non-specific	1	156	3.67e-14	pfam01121	CoaE	cl17190	Dephospho-CoA kinase
	non-specific	2	156	6.80e-06	PRK14732	coaE	cl17190	Dephospho-CoA kinase
	non-specific	1	156	8.58e-05	PRK14734	coaE	cl17190	Dephospho-CoA kinase
	non-specific	2	185	0.000155	PRK14731	coaE	cl17190	Dephospho-CoA kinase
	non-specific	3	153	0.000692	pfam13521	AAA_28	cl21455	AAA domain
	superfamily	3	153	0.000692	cl21455	P-loop_NTPase superfamily	-	P-loop containing Nucleoside Triphosphate Hydrolases
	non-specific	1	153	0.001474	COG1428	Dck	cl17190	Deoxyadenosine/deoxycytidine kinase
	non-specific	5	90	0.003355	PLN02422	PLN02422	cl17190	Dephospho-CoA kinase
	non-specific	1	28	0.003534	PRK04182	PRK04182	cl28332	cytidylate kinase
	superfamily	1	28	0.003534	cl28332	CmkB superfamily	-	Cytidylate kinase
	non-specific	3	35	0.006008	cd01983	Fer4_NifH	cl28886	The Fer4_NifH superfamily
	superfamily	3	35	0.006008	cl28886	Fer4_NifH superfamily	-	The Fer4_NifH superfamily
	non-specific	1	37	0.009106	COG1102	CmkB	cl28332	Cytidylate kinase
M. primatum	non-specific	1	144	2.19e-18	COG0237	CoaE	cl28605	Dephospho-CoA kinase
	superfamily	1	144	2.19e-18	cl28605	CoaE superfamily	-	Dephospho-CoA kinase
	non-specific	1	148	2.85e-15	cd02022	DPCK	cl17190	Dephospho-CoA kinase
	superfamily	1	148	2.85e-15	cl17190	NK superfamily	-	Nucleoside/nucleotide kinase superfamily
	non-specific	2	137	4.48e-10	TIGR00152	TIGR00152	cl17190	Dephospho-CoA kinase
	non-specific	2	148	2.85e-05	PRK00081	coaE	cl17190	Dephospho-CoA kinase
M. pulmonis	specific	1	172	1.38e-34	COG0237	CoaE	cl28605	Dephospho-CoA kinase
	superfamily	1	172	1.38e-34	cl28605	CoaE superfamily	-	Dephospho-CoA kinase
	specific	1	170	1.51e-29	cd02022	DPCK	cl17190	Dephospho-CoA kinase
	superfamily	1	170	1.51e-29	cl17190	NK superfamily	-	Nucleoside/nucleotide kinase superfamily

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. pulmonis	non-specific	2	137	2.64e-20	TIGR00152	TIGR00152	cl17190	Dephospho-CoA kinase
	non-specific	2	136	6.79e-12	PRK00081	coaE	cl17190	Dephospho-CoA kinase
	non-specific	2	143	9.80e-06	pfam01121	CoaE	cl17190	Dephospho-CoA kinase
	non-specific	2	143	3.27e-05	pfam13238	AAA_18	cl21455	AAA domain
	superfamily	2	143	3.27e-05	cl21455	P-loop_NTPase superfamily	-	P-loop containing Nucleoside Triphosphate Hydrolases
	non-specific	1	78	0.000167	COG1936	Fap7	cl17190	Broad-specificity NMP kinase
	non-specific	2	27	0.0002	cd02034	CooC	cl28886	The accessory protein CooC, which contains a nucleotide-binding domain (P-loop) near the N-terminus, participates in the maturation of the nickel center of carbon monoxide dehydrogenase (CODH).
	superfamily	2	27	0.0002	cl28886	Fer4_NifH superfamily	-	The Fer4_NifH superfamily
	non-specific	2	140	0.001408	PRK14733	coaE	cl17190	Dephospho-CoA kinase
	non-specific	2	171	0.002499	COG1100	Gem1	cl27030	GTPase SAR1 family domain [General function prediction only]
	superfamily	2	171	0.002499	cl27030	Ras superfamily	-	Ras family
	non-specific	3	157	0.002851	TIGR01054	rgy	cl27598	reverse gyrase
	superfamily	3	157	0.002851	cl27598	TOP1Bc superfamily	-	Bacterial DNA topoisomeraes I ATP-binding domain
	non-specific	2	146	0.003434	COG1102	CmkB	cl28332	Cytidylate kinase
	superfamily	2	146	0.003434	cl28332	CmkB superfamily	-	Cytidylate kinase
	non-specific	2	30	0.004023	COG3640	CooC	cl27521	CO dehydrogenase nickel-insertion accessory protein CooC1 [Posttranslational modification, protein turnover, chaperones]
	superfamily	2	30	0.004023	cl27521	CbiA superfamily	-	CobQ/CobB/MinD/ParA nucleotide binding domain
	non-specific	2	27	0.005486	cd01983	Fer4_NifH	cl28886	The Fer4_NifH superfamily
M. putrefaciens	specific	2	168	2.03e-25	cd02022	DPCK	cl17190	Dephospho-CoA kinase
	superfamily	2	168	2.03e-25	cl17190	NK superfamily	-	Nucleoside/nucleotide kinase superfamily
	non-specific	1	144	6.61e-20	PRK00081	coaE	cl17190	Dephospho-CoA kinase
	non-specific	1	144	1.37e-19	pfam01121	CoaE	cl17190	Dephospho-CoA kinase
	non-specific	1	144	7.97e-17	COG0237	CoaE	cl28605	Dephospho-CoA kinase
	superfamily	1	144	7.97e-17	cl28605	CoaE superfamily	-	Dephospho-CoA kinase
	non-specific	2	162	1.78e-15	TIGR00152	TIGR00152	cl17190	Dephospho-CoA kinase
	non-specific	2	145	1.48e-08	PRK14730	coaE	cl17190	Dephospho-CoA kinase
	non-specific	2	166	8.52e-08	PRK14731	coaE	cl17190	Dephospho-CoA kinase
	non-specific	3	141	2.13e-06	PRK14733	coaE	cl17190	Dephospho-CoA kinase
	non-specific	3	93	2.83e-05	PRK14734	coaE	cl17190	Dephospho-CoA kinase
	non-specific	2	163	0.000238	PRK14732	coaE	cl17190	Dephospho-CoA kinase
	non-specific	1	150	0.000584	TIGR02173	cyt_kin_arch	cl28332	cytidylate kinase
	superfamily	1	150	0.000584	cl28332	CmkB superfamily	-	Cytidylate kinase
	non-specific	1	56	0.000825	COG1102	CmkB	cl28332	Cytidylate kinase
	non-specific	2	94	0.003407	PLN02422	PLN02422	cl17190	Dephospho-CoA kinase

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. putrefaciens	non-specific	10	86	0.003775	pfam13207	AAA_17	cl21455	AAA domain
	superfamily	10	86	0.003775	cl21455	P-loop_NTPase superfamily	-	P-loop containing Nucleoside Triphosphate Hydrolases
	non-specific	2	141	0.005687	cd02023	UMPK	cl17190	Uridine monophosphate kinase
	non-specific	2	141	0.006565	TIGR00235	udk	cl17190	Uridine kinase superfamily
M. simbae	non-specific	1	137	5.63e-19	COG0237	CoaE	cl28605	Dephospho-CoA kinase
	superfamily	1	137	5.63e-19	cl28605	CoaE superfamily	-	Dephospho-CoA kinase
	non-specific	1	137	2.29e-16	cd02022	DPCK	cl17190	Dephospho-CoA kinase
	superfamily	1	137	2.29e-16	cl17190	NK superfamily	-	Nucleoside/nucleotide kinase superfamily
	non-specific	2	146	5.49e-13	TIGR00152	TIGR00152	cl17190	Dephospho-CoA kinase
	non-specific	2	137	1.57e-09	PRK00081	coaE	cl17190	Dephospho-CoA kinase
	non-specific	2	137	5.44e-05	pfam01121	CoaE	cl17190	Dephospho-CoA kinase
	non-specific	2	51	0.002071	TIGR03375	type_I_sec_LssB	cl26602	type I secretion system ATPase, LssB family
	superfamily	2	51	0.002071	cl26602	SunT superfamily	-	ABC-type bacteriocin/lantibiotic exporters, contain an N-terminal double-glycine peptidase domain [Defense mechanisms]
	non-specific	1	17	0.007807	COG0488	Uup	cl28181	ATPase components of ABC transporters with duplicated ATPase domains [General function prediction only]
	superfamily	1	17	0.007807	cl28181	AAA superfamily	-	ATPases associated with a variety of cellular activities
M. sturni	non-specific	1	187	8.24e-10	COG0237	CoaE	cl28605	Dephospho-CoA kinase
	superfamily	1	187	8.24e-10	cl28605	CoaE superfamily	-	Dephospho-CoA kinase
	non-specific	1	129	6.98e-09	cd02022	DPCK	cl17190	Dephospho-CoA kinase
	superfamily	1	129	6.98e-09	cl17190	NK superfamily	-	Nucleoside/nucleotide kinase superfamily
	non-specific	2	110	5.18e-05	PRK00081	coaE	cl17190	Dephospho-CoA kinase
	non-specific	2	128	0.000306	TIGR00152	TIGR00152	cl17190	Dephospho-CoA kinase
M. synoviae	non-specific	1	141	1.74e-12	COG0237	CoaE	cl28605	Dephospho-CoA kinase
	superfamily	1	141	1.74e-12	cl28605	CoaE superfamily	-	Dephospho-CoA kinase
	non-specific	1	147	2.92e-11	cd02022	DPCK	cl17190	Dephospho-CoA kinase
	superfamily	1	147	2.92e-11	cl17190	NK superfamily	-	Nucleoside/nucleotide kinase superfamily
	non-specific	2	166	1.06e-06	TIGR00152	TIGR00152	cl17190	Dephospho-CoA kinase
	non-specific	1	76	0.006477	pfam05272	VirE	cl23993	Virulence-associated protein E (these proteins contain a P-loop motif)
	superfamily	1	76	0.006477	cl23993	VirE superfamily	-	Virulence-associated protein E (these proteins contain a P-loop motif)
M. testudinis	specific	2	175	2.12e-33	cd02022	DPCK	cl17190	Dephospho-CoA kinase
	superfamily	2	175	2.12e-33	cl17190	NK superfamily	-	Nucleoside/nucleotide kinase superfamily
	non-specific	5	188	4.62e-32	PRK00081	coaE	cl17190	Dephospho-CoA kinase
	specific	1	193	1.37e-31	COG0237	CoaE	cl28605	Dephospho-CoA kinase
	superfamily	1	193	1.37e-31	cl28605	CoaE superfamily	-	Dephospho-CoA kinase
	non-specific	2	143	9.59e-29	TIGR00152	TIGR00152	cl17190	Dephospho-CoA kinase
	non-specific	1	89	3.48e-16	pfam01121	CoaE	cl17190	Dephospho-CoA kinase

Query	Hit type	ID region start	ID region end	E-Value	Accession	Short name	Superfamily	Definition
M. testudinis	non-specific	2	191	6.50e-13	PRK14732	coaE	cl17190	Dephospho-CoA kinase
	non-specific	5	175	1.07e-11	PLN02422	PLN02422	cl17190	Dephospho-CoA kinase
	non-specific	2	187	4.17e-11	PRK14731	coaE	cl17190	Dephospho-CoA kinase
	non-specific	7	89	8.21e-08	PRK14730	coaE	cl17190	Dephospho-CoA kinase
	non-specific	1	89	3.35e-06	PRK03333	coaE	cl28605	Dephospho-CoA kinase
	non-specific	1	89	0.00045	PRK14734	coaE	cl17190	Dephospho-CoA kinase
	non-specific	2	15	0.001737	PRK00349	uvrA	cl26603	excinuclease ABC subunit A
	superfamily	2	15	0.001737	cl26603	UvrA superfamily	-	Excinuclease UvrABC ATPase subunit [Replication, recombination and repair]
	non-specific	1	175	0.002035	PTZ00451	PTZ00451	cl17190	Dephospho-CoA kinase
	non-specific	2	19	0.003434	PRK05416	PRK05416	cl23728	glmZ(sRNA)-inactivating NTPase
	superfamily	2	19	0.003434	cl23728	ATP_bind_2 superfamily	-	P-loop ATPase protein family
	non-specific	2	21	0.007101	COG0178	UvrA	cl26603	Excinuclease UvrABC ATPase subunit [Replication, recombination and repair]
M. yeatsii	specific	2	169	2.21e-25	cd02022	DPCK	cl17190	Dephospho-CoA kinase
	superfamily	2	169	2.21e-25	cl17190	NK superfamily	-	Nucleoside/nucleotide kinase superfamily
	non-specific	1	145	5.32e-18	PRK00081	coaE	cl17190	Dephospho-CoA kinase
	non-specific	1	99	6.31e-18	pfam01121	CoaE	cl17190	Dephospho-CoA kinase
	non-specific	1	176	6.42e-17	COG0237	CoaE	cl28605	Dephospho-CoA kinase
	superfamily	1	176	6.42e-17	cl28605	CoaE superfamily	-	Dephospho-CoA kinase
	non-specific	2	163	9.11e-14	TIGR00152	TIGR00152	cl17190	Dephospho-CoA kinase
	non-specific	3	93	1.16e-08	PRK14734	coaE	cl17190	Dephospho-CoA kinase
	non-specific	2	99	3.34e-08	PRK14730	coaE	cl17190	Dephospho-CoA kinase
	non-specific	2	98	2.12e-07	PLN02422	PLN02422	cl17190	Dephospho-CoA kinase
	non-specific	3	143	4.58e-07	PRK14733	coaE	cl17190	Dephospho-CoA kinase
	non-specific	2	183	3.61e-06	PRK14732	coaE	cl17190	Dephospho-CoA kinase
	non-specific	3	77	0.004789	PRK03333	coaE	cl28605	Dephospho-CoA kinase

Supplementary Table 1.7 The InterPro search results of the PanK type III amino acid sequences of the currently annotated CoA biosynthetic pathway enzyme-encoding genes and newly identified homologues

Mycoplasma species	Amino acid region	Database	Database ID	Database signature description	ID region start	ID region end	InterPro ID[b]	ID type[a]	Gene Ontology (GO) term[c]
M. alligatoris	253	Gene3D	G3DSA:3.30.420.40		4	134			
		PANTHER	PTHR34265		9	251	IPR004619	F	GO:0004594
		Pfam	PF03309	Type III pantothenate kinase	8	193	IPR004619	F	GO:0004594
		SUPERFAMILY	SSF53067		123	244			
		SUPERFAMILY	SSF53067		8	116			
		TIGRFAM	TIGR00671	baf: pantothenate kinase, type III	8	224	IPR004619	F	GO:0004594
M. alvi	257	CDD	cd00012	NBD_sugar-kinase_HSP70_actin	9	156			
		Gene3D	G3DSA:3.30.420.40		1	84			
		Gene3D	G3DSA:3.30.420.40		102	257			
		Hamap	MF_01274	Type III pantothenate kinase [coaX].	7	256	IPR004619		GO:0004594
		PANTHER	PTHR34265		8	255	IPR004619	F	GO:0004594
		Pfam	PF03309	Type III pantothenate kinase	9	202	IPR004619	F	GO:0004594
		SUPERFAMILY	SSF53067		9	121			
		SUPERFAMILY	SSF53067		127	252			
		TIGRFAM	TIGR00671	baf: pantothenate kinase, type III	9	228	IPR004619	F	GO:0004594
M. anatis	268	CDD	cd00012	NBD_sugar-kinase_HSP70_actin	8	172			
		Gene3D	G3DSA:3.30.420.40		3	143			
		Pfam	PF03309	Type III pantothenate kinase	8	173	IPR004619	F	GO:0004594
		SUPERFAMILY	SSF53067		7	114			
M. arginini	250	Gene3D	G3DSA:3.30.420.40		94	250			
		Hamap	MF_01274	Type III pantothenate kinase [coaX].	8	250	IPR004619	F	GO:0004594
		PANTHER	PTHR34265		10	249	IPR004619	F	GO:0004594
		Pfam	PF03309	Type III pantothenate kinase	10	206	IPR004619	F	GO:0004594
		SUPERFAMILY	SSF53067		126	249			
		SUPERFAMILY	SSF53067		9	120			
		TIGRFAM	TIGR00671	baf: pantothenate kinase, type III	10	234	IPR004619	F	GO:0004594
M. buteonis	246	Gene3D	G3DSA:3.30.420.40		1	100			
		Gene3D	G3DSA:3.30.420.40		101	240			
		Pfam	PF03309	Type III pantothenate kinase	5	198	IPR004619	F	GO:0004594
M. columborale	273	Gene3D	G3DSA:3.30.420.40		104	273			
		PANTHER	PTHR34265		5	271	IPR004619	F	GO:0004594
		Pfam	PF03309	Type III pantothenate kinase	5	209	IPR004619	F	GO:0004594
		SUPERFAMILY	SSF53067		129	235			
M. cricetuli	272	Gene3D	G3DSA:3.30.420.40		97	272			
		PANTHER	PTHR34265		4	242	IPR004619	F	GO:0004594
		Pfam	PF03309	Type III pantothenate kinase	4	209	IPR004619	F	GO:0004594
		SUPERFAMILY	SSF53067		133	242			
		SUPERFAMILY	SSF53067		4	124			
M. crocodyli	258	Gene3D	G3DSA:3.30.420.40		108	258			
		Hamap	MF_01274	Type III pantothenate kinase [coaX].	12	257	IPR004619	F	GO:0004594

Mycoplasma species	Amino acid region	Database	Database ID	Database signature description	ID region start	ID region end	InterPro ID[b]	ID type[a]	Gene Ontology (GO) term[c]
M. crocodyli	258	PANTHER	PTHR34265		14	256	IPR004619	F	GO:0004594
		Pfam	PF03309	Type III pantothenate kinase	13	203	IPR004619	F	GO:0004594
		SUPERFAMILY	SSF53067		128	250			
		SUPERFAMILY	SSF53067		13	123			
		TIGRFAM	TIGR00671	baf: pantothenate kinase, type III	14	241	IPR004619	F	GO:0004594
M. gallinaceum	248	Gene3D	G3DSA:3.30.420.40		90	247			
		PANTHER	PTHR34265		6	245	IPR004619	F	GO:0004594
		Pfam	PF03309	Type III pantothenate kinase	6	196	IPR004619	F	GO:0004594
		SUPERFAMILY	SSF53067		6	113			
		SUPERFAMILY	SSF53067		119	245			
		TIGRFAM	TIGR00671	baf: pantothenate kinase, type III	7	206	IPR004619	F	GO:0004594
M. iowae	258	Gene3D	G3DSA:3.30.420.40		89	258			
		Hamap	MF_01274	Type III pantothenate kinase [coaX].	4	252	IPR004619	F	GO:0004594
		PANTHER	PTHR34265		5	251	IPR004619	F	GO:0004594
		Pfam	PF03309	Type III pantothenate kinase	5	204	IPR004619	F	GO:0004594
		SUPERFAMILY	SSF53067		5	114			
		SUPERFAMILY	SSF53067		123	252			
		TIGRFAM	TIGR00671	baf: pantothenate kinase, type III	6	236	IPR004619	F	GO:0004594
M. mobile	243	Gene3D	G3DSA:3.30.420.40		88	242			
		Hamap	MF_01274	Type III pantothenate kinase [coaX].	1	243	IPR004619	F	GO:0004594
		PANTHER	PTHR34265		1	236	IPR004619	F	GO:0004594
		Pfam	PF03309	Type III pantothenate kinase	3	197	IPR004619	F	GO:0004594
		SUPERFAMILY	SSF53067		1	115			
		SUPERFAMILY	SSF53067		117	240			
		TIGRFAM	TIGR00671	baf: pantothenate kinase, type III	3	235	IPR004619	F	GO:0004594
M. molare	242	Gene3D	G3DSA:3.30.420.40		90	238			
		Hamap	MF_01274	Type III pantothenate kinase [coaX].	1	242	IPR004619	F	GO:0004594
		PANTHER	PTHR34265		2	236	IPR004619	F	GO:0004594
		Pfam	PF03309	Type III pantothenate kinase	1	198	IPR004619	F	GO:0004594
		SUPERFAMILY	SSF53067		117	228			
		SUPERFAMILY	SSF53067		1	115			
		TIGRFAM	TIGR00671	baf: pantothenate kinase, type III	1	228	IPR004619	F	GO:0004594
M. penetrans	255	Gene3D	G3DSA:3.30.420.40		4	138			
		Hamap	MF_01274	Type III pantothenate kinase [coaX].	7	255	IPR004619	F	GO:0004594
		PANTHER	PTHR34265		8	253	IPR004619	F	GO:0004594
		Pfam	PF03309	Type III pantothenate kinase	8	203	IPR004619	F	GO:0004594
		SUPERFAMILY	SSF53067		124	253			
		SUPERFAMILY	SSF53067		7	119			
		TIGRFAM	TIGR00671	baf: pantothenate kinase, type III	9	229	IPR004619	F	GO:0004594
M. pirum	256	Gene3D	G3DSA:3.30.420.40		93	255			
		PANTHER	PTHR34265		8	248	IPR004619	F	GO:0004594
		Pfam	PF03309	Type III pantothenate kinase	9	204	IPR004619	F	GO:0004594
		SUPERFAMILY	SSF53067		125	254			

Mycoplasma species	Amino acid region	Database	Database ID	Database signature description	ID region start	ID region end	InterPro ID[b]	ID type[a]	Gene Ontology (GO) term[c]
M. pirum	256	SUPERFAMILY	SSF53067		9	120			
		TIGRFAM	TIGR00671	baf: pantothenate kinase, type III	9	231	IPR004619	F	GO:0004594
M. pulmonis	224	Gene3D	G3DSA:3.30.420.40		96	222			
		PANTHER	PTHR34265		1	217	IPR004619	F	GO:0004594
		Pfam	PF03309	Type III pantothenate kinase	3	203	IPR004619	F	GO:0004594
		SUPERFAMILY	SSF53067		121	218			
		SUPERFAMILY	SSF53067		1	118			
		TIGRFAM	TIGR00671	baf: pantothenate kinase, type III	4	217	IPR004619	F	GO:0004594
M. sturni	273	Gene3D	G3DSA:3.30.420.40		100	273			
		Pfam	PF03309	Type III pantothenate kinase	6	210	IPR004619	F	GO:0004594
		Phobius	CYTOPLASMIC_DOMAIN	Region of a membrane-bound protein predicted to be outside the membrane, in the cytoplasm.	217	273			
		Phobius	TRANSMEMBRANE	Region of a membrane-bound protein predicted to be embedded in the membrane.	196	216			
		Phobius	NON_CYTOPLASMIC_DOMAIN	Region of a membrane-bound protein predicted to be outside the membrane, in the extracellular region.	1	195			
		SUPERFAMILY	SSF53067		130	237			
		TMHMM	TMhelix	Region of a membrane-bound protein predicted to be embedded in the membrane.	194	216			
M. synoviae	250	Gene3D	G3DSA:3.30.420.40		94	250			
		Hamap	MF_01274	Type III pantothenate kinase [coaX].	8	250	IPR004619	F	GO:0004594
		PANTHER	PTHR34265		10	249	IPR004619	F	GO:0004594
		Pfam	PF03309	Type III pantothenate kinase	10	206	IPR004619	F	GO:0004594
		SUPERFAMILY	SSF53067		9	120			
		SUPERFAMILY	SSF53067		126	249			
		TIGRFAM	TIGR00671	baf: pantothenate kinase, type III	10	234	IPR004619	F	GO:0004594
M. testudinis	248	Gene3D	G3DSA:3.30.420.40		92	246			
		Hamap	MF_01274	Type III pantothenate kinase [coaX].	3	244	IPR004619	F	GO:0004594
		PANTHER	PTHR34265		5	242	IPR004619	F	GO:0004594
		Pfam	PF03309	Type III pantothenate kinase	5	196	IPR004619	F	GO:0004594
		SUPERFAMILY	SSF53067		4	119			
		SUPERFAMILY	SSF53067		121	240			
		TIGRFAM	TIGR00671	baf: pantothenate kinase, type III	5	240	IPR004619	F	GO:0004594

[a]ID type abbreviations – F, Family
[b]InterPro ID – IPR004619: Type III pantothenate kinase
[c]GO term (Molecular Function) – GO:0004594: pantothenate kinase activity

Supplementary Table 1.8 The InterPro search results of the CoaBC amino acid sequences of the currently annotated CoA biosynthetic pathway enzyme-encoding genes and newly identified homologues

Mycoplasma species	Amino acid region	Database	Database ID	Database signature description	ID region start	ID region end	InterPro ID[b]	ID type[a]	Gene Ontology (GO) term[c]
M. anatis	375	Gene3D	G3DSA:3.40.50.1950		1	165	IPR036551	H	GO:0003824
		Gene3D	G3DSA:3.40.50.10300		166	375	IPR035929	H	
		PANTHER	PTHR14359:SF6		2	334			
		PANTHER	PTHR14359		2	334			
		Pfam	PF02441	Flavoprotein	1	163	IPR003382	D	GO:0003824
		Pfam	PF04127	DNA / pantothenate metabolism flavoprotein	170	335	IPR007085	D	
		Phobius	SIGNAL_PEPTIDE_N_REGION	N-terminal region of a signal peptide.	1	2			
		Phobius	SIGNAL_PEPTIDE_H_REGION	Hydrophobic region of a signal peptide.	3	13			
		Phobius	SIGNAL_PEPTIDE_C_REGION	C-terminal region of a signal peptide.	14	18			
		Phobius	SIGNAL_PEPTIDE	Signal peptide region	1	18			
		Phobius	NON_CYTOPLASMIC_D	Region of a membrane-bound protein predicted to be outside the membrane, in the extracellular region.	19	375			
		SUPERFAMILY	SSF52507		1	165	IPR036551	H	GO:0003824
		SUPERFAMILY	SSF102645		170	373	IPR035929	H	
		TIGRFAM	TIGR00521	coaBC_dfp: phosphopantothenoylcysteine decarboxylase / phosphopantothenate--cysteine ligase	1	371	IPR005252	F	GO:0004632; GO:0004633; GO:0010181; GO:0015937; GO:0015941
M. arginini	384	Gene3D	G3DSA:3.40.50.1950		1	168	IPR036551	H	GO:0003824
		Gene3D	G3DSA:3.40.50.10300		169	384	IPR035929	H	
		PANTHER	PTHR14359		2	344			
		PANTHER	PTHR14359:SF6		2	344			
		Pfam	PF04127	DNA / pantothenate metabolism flavoprotein	176	342	IPR007085	D	
		Pfam	PF02441	Flavoprotein	1	171	IPR003382	D	GO:0003824
		Phobius	NON_CYTOPLASMIC_D	Region of a membrane-bound protein predicted to be outside the membrane, in the extracellular region.	17	384			
		Phobius	SIGNAL_PEPTIDE_C_REGION	C-terminal region of a signal peptide.	12	16			
		Phobius	SIGNAL_PEPTIDE	Signal peptide region	1	16			
		Phobius	SIGNAL_PEPTIDE_N_REGION	N-terminal region of a signal peptide.	1	2			
		Phobius	SIGNAL_PEPTIDE_H_REGION	Hydrophobic region of a signal peptide.	3	11			
		SignalP_EUK	SignalP-noTM		1	17			
		SUPERFAMILY	SSF102645		176	381	IPR035929	H	
		SUPERFAMILY	SSF52507		1	169	IPR036551	H	GO:0003824
		TIGRFAM	TIGR00521	coaBC_dfp: phosphopantothenoylcysteine decarboxylase / phosphopantothenate--cysteine ligase	1	376	IPR005252	F	GO:0004632; GO:0004633; GO:0010181; GO:0015937; GO:0015941

Mycoplasma species	Amino acid region	Database	Database ID	Database signature description	ID region start	ID region end	InterPro ID[b]	ID type[a]	Gene Ontology (GO) term[c]
M. columborale	377	Gene3D	G3DSA:3.40.50.10300		165	376	IPR035929	H	
		Gene3D	G3DSA:3.40.50.1950		1	164	IPR036551	H	GO:0003824
		PANTHER	PTHR14359:SF6		2	341			
		PANTHER	PTHR14359		2	341			
		Pfam	PF04127	DNA / pantothenate metabolism flavoprotein	171	335	IPR007085	D	
		Pfam	PF02441	Flavoprotein	1	167	IPR003382	D	GO:0003824
		Phobius	SIGNAL_PEPTIDE	Signal peptide region	1	16			
		Phobius	SIGNAL_PEPTIDE_C_REGION	C-terminal region of a signal peptide.	12	16			
		Phobius	SIGNAL_PEPTIDE_N_REGION	N-terminal region of a signal peptide.	1	2			
		Phobius	SIGNAL_PEPTIDE_H_REGION	Hydrophobic region of a signal peptide.	3	11			
		Phobius	NON_CYTOPLASMIC_D	Region of a membrane-bound protein predicted to be outside the membrane, in the extracellular region.	17	377			
		SUPERFAMILY	SSF52507		1	166	IPR036551	H	GO:0003824
		SUPERFAMILY	SSF102645		170	373	IPR035929	H	
		TIGRFAM	TIGR00521	coaBC_dfp: phosphopantothenoylcysteine decarboxylase / phosphopantothenate--cysteine ligase	1	370	IPR005252	F	GO:0004632; GO:0004633; GO:0010181; GO:0015937; GO:0015941
M. cricetuli	377	Gene3D	G3DSA:3.40.50.10300		169	377	IPR035929	H	
		Gene3D	G3DSA:3.40.50.1950		1	168	IPR036551	H	GO:0003824
		PANTHER	PTHR14359		2	344			
		PANTHER	PTHR14359:SF6		2	344			
		Pfam	PF02441	Flavoprotein	1	167	IPR003382	D	GO:0003824
		Pfam	PF04127	DNA / pantothenate metabolism flavoprotein	169	336	IPR007085	D	
		Phobius	SIGNAL_PEPTIDE	Signal peptide region	1	16			
		Phobius	SIGNAL_PEPTIDE_C_REGION	C-terminal region of a signal peptide.	12	16			
		Phobius	NON_CYTOPLASMIC_D	Region of a membrane-bound protein predicted to be outside the membrane, in the extracellular region.	17	377			
		Phobius	SIGNAL_PEPTIDE_H_REGION	Hydrophobic region of a signal peptide.	3	11			
		Phobius	SIGNAL_PEPTIDE_N_REGION	N-terminal region of a signal peptide.	1	2			
		SUPERFAMILY	SSF102645		169	376	IPR035929	H	
		SUPERFAMILY	SSF52507		1	167	IPR036551	H	GO:0003824
		TIGRFAM	TIGR00521	coaBC_dfp: phosphopantothenoylcysteine decarboxylase / phosphopantothenate--cysteine ligase	1	371	IPR005252	F	GO:0004632; GO:0004633; GO:0010181; GO:0015937; GO:0015941
M. gallinaceum	385	Gene3D	G3DSA:3.40.50.1950		1	176	IPR036551	H	GO:0003824
		Gene3D	G3DSA:3.40.50.10300		177	385	IPR035929	H	
		PANTHER	PTHR14359		2	342			
		PANTHER	PTHR14359:SF6		2	342			

Mycoplasma species	Amino acid region	Database	Database ID	Database signature description	ID region start	ID region end	InterPro ID[b]	ID type[a]	Gene Ontology (GO) term[c]
M. gallinaceum	385	Pfam	PF04127	DNA / pantothenate metabolism flavoprotein	178	348	IPR007085	D	
		Pfam	PF02441	Flavoprotein	1	173	IPR003382	D	GO:0003824
		SUPERFAMILY	SSF52507		1	173	IPR036551	H	GO:0003824
		SUPERFAMILY	SSF102645		178	382	IPR035929	H	
		TIGRFAM	TIGR00521	coaBC_dfp: phosphopantothenoylcysteine decarboxylase / phosphopantothenate--cysteine ligase	1	377	IPR005252	F	GO:0004632; GO:0004633; GO:0010181; GO:0015937; GO:0015941
M. iowae (PPCS)	236	Gene3D	G3DSA:3.40.50.10300		1	236	IPR035929	H	
		PANTHER	PTHR14359:SF22		1	226			
		PANTHER	PTHR14359		1	226			
		Pfam	PF04127	DNA / pantothenate metabolism flavoprotein	2	195	IPR007085	D	
		SUPERFAMILY	SSF102645		1	232	IPR035929	H	
M. iowae (PPCDC)	178	Gene3D	G3DSA:3.40.50.1950		1	177	IPR036551	H	GO:0003824
		PANTHER	PTHR14359		2	174			
		PANTHER	PTHR14359:SF6		2	174			
		Pfam	PF02441	Flavoprotein	3	175	IPR003382	D	GO:0003824
		ProSiteProfiles	PS51257	Prokaryotic membrane lipoprotein lipid attachment site profile.	1	15			
		SUPERFAMILY	SSF52507		2	175	IPR036551	H	GO:0003824
M. mobile	360	Gene3D	G3DSA:3.40.50.1950		1	171	IPR036551	H	GO:0003824
		Gene3D	G3DSA:3.40.50.10300		172	230	IPR035929	H	
		Gene3D	G3DSA:3.40.50.10300		231	360	IPR035929	H	
		PANTHER	PTHR14359		3	323			
		PANTHER	PTHR14359:SF6		3	323			
		Pfam	PF02441	Flavoprotein	2	164	IPR003382	D	GO:0003824
		Pfam	PF04127	DNA / pantothenate metabolism flavoprotein	175	320	IPR007085	D	
		SUPERFAMILY	SSF102645		172	358	IPR035929	H	
		SUPERFAMILY	SSF52507		2	170	IPR036551	H	GO:0003824
M. sturni	376	Gene3D	G3DSA:3.40.50.1950		1	165	IPR036551	H	GO:0003824
		Gene3D	G3DSA:3.40.50.10300		166	376	IPR035929	H	
		PANTHER	PTHR14359:SF6		2	334			
		PANTHER	PTHR14359		2	334			
		Pfam	PF02441	Flavoprotein	1	164	IPR003382	D	GO:0003824
		Pfam	PF04127	DNA / pantothenate metabolism flavoprotein	170	334	IPR007085	D	
		Phobius	SIGNAL_PEPTIDE_N_REGION	N-terminal region of a signal peptide.	1	2			
		Phobius	SIGNAL_PEPTIDE_C_REGION	C-terminal region of a signal peptide.	15	18			
		Phobius	SIGNAL_PEPTIDE	Signal peptide region	1	18			
		Phobius	NON_CYTOPLASMIC_D	Region of a membrane-bound protein predicted to be outside the membrane, in the extracellular region.	19	376			
		Phobius	SIGNAL_PEPTIDE_H_REGION	Hydrophobic region of a signal peptide.	3	14			

Mycoplasma species	Amino acid region	Database	Database ID	Database signature description	ID region start	ID region end	InterPro ID[b]	ID type[a]	Gene Ontology (GO) term[c]
M. sturni	376	SUPERFAMILY	SSF102645		170	375	IPR035929	H	
		SUPERFAMILY	SSF52507		1	164	IPR036551	H	GO:0003824
		TIGRFAM	TIGR00521	coaBC_dfp: phosphopantothenoylcysteine decarboxylase / phosphopantothenate--cysteine ligase	2	370	IPR005252	F	GO:0004632; GO:0004633; GO:0010181; GO:0015937; GO:0015941
M. synoviae	384	Gene3D	G3DSA:3.40.50.10300		169	384	IPR035929	H	
		Gene3D	G3DSA:3.40.50.1950		1	168	IPR036551	H	GO:0003824
		PANTHER	PTHR14359:SF6		2	344			
		PANTHER	PTHR14359		2	344			
		Pfam	PF04127	DNA / pantothenate metabolism flavoprotein	176	342	IPR007085	D	
		Pfam	PF02441	Flavoprotein	1	171	IPR003382	D	GO:0003824
		Phobius	SIGNAL_PEPTIDE_H_REGION	Hydrophobic region of a signal peptide.	3	11			
		Phobius	SIGNAL_PEPTIDE	Signal peptide region	1	16			
		Phobius	SIGNAL_PEPTIDE_N_REGION	N-terminal region of a signal peptide.	1	2			
		Phobius	NON_CYTOPLASMIC_D	Region of a membrane-bound protein predicted to be outside the membrane, in the extracellular region.	17	384			
		Phobius	SIGNAL_PEPTIDE_C_REGION	C-terminal region of a signal peptide.	12	16			
		SignalP_EUK	SignalP-noTM		1	17			
		SUPERFAMILY	SSF102645		176	381	IPR035929	H	
		SUPERFAMILY	SSF52507		1	169	IPR036551	H	GO:0003824
		TIGRFAM	TIGR00521	coaBC_dfp: phosphopantothenoylcysteine decarboxylase / phosphopantothenate--cysteine ligase	1	376	IPR005252	F	GO:0004632; GO:0004633; GO:0010181; GO:0015937; GO:0015941
M. testudinis	380	Gene3D	G3DSA:3.40.50.1950		13	195	IPR036551	H	GO:0003824
		Gene3D	G3DSA:3.40.50.10300		196	246	IPR035929	H	
		Gene3D	G3DSA:3.40.50.10300		247	380	IPR035929	H	
		PANTHER	PTHR14359		17	242			
		Pfam	PF02441	Flavoprotein	17	186	IPR003382	D	GO:0003824
		Pfam	PF04127	DNA / pantothenate metabolism flavoprotein	196	347	IPR007085	D	
		SUPERFAMILY	SSF52507		17	190	IPR036551	H	GO:0003824
		SUPERFAMILY	SSF102645		196	379	IPR035929	H	

[a]ID type abbreviations – H, Homologous superfamily; F, Family; D, Domain

[b]InterPro ID – IPR036551: Flavin prenyltransferase-like; IPR035929: CoaB-like superfamily; IPR005252: Coenzyme A biosynthesis bifunctional protein, CoaBC; IPR003382: Flavoprotein; IPR007085: DNA/pantothenate metabolism flavoprotein, C-terminal

[c]GO term (Biological Process) – GO:0015937: Coenzyme A biosynthetic process; GO:0015941: Pantothenate catabolic process
GO term (Molecular Function) – GO:0003824: catalytic activity; GO:0004632: Phosphopantothenate--cysteine ligase activity; GO:0004633: Phosphopantothenoylcysteine decarboxylase activity; GO:0010181: FMN binding

Supplementary Table 1.9 The InterPro search results of the PPAT amino acid sequences of the currently annotated CoA biosynthetic pathway enzyme-encoding genes and newly identified homologues

Mycoplasma species	Amino acid region	Database	Database ID	Database signature description	ID region start	ID region end	InterPro ID	ID type	Gene Ontology (GO) term
M. agalactiae	140	Gene3D	G3DSA:3.40.50.620		1	140	IPR014729	H	
		Hamap	MF_00151	Phosphopantetheine adenylyltransferase [coaD].	2	140	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342:SF1		1	139	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342		1	139			
		Pfam	PF01467	Cytidylyltransferase-like	5	134	IPR004821	D	GO:0003824; GO:0009058
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	2	20	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	49	73	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	20	41	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	113	135	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	86	102	IPR001980	F	GO:0004595; GO:0015937
		SUPERFAMILY	SSF52374		2	139			
		TIGRFAM	TIGR01510	coaD_prev_kdtB: pantetheine-phosphate adenylyltransferase	4	134	IPR001980	F	GO:0004595; GO:0015937
		TIGRFAM	TIGR00125	cyt_tran_rel: cytidyltransferase-like domain	4	58	IPR004821	D	GO:0003824; GO:0009058
M. alligatoris	145	Gene3D	G3DSA:3.40.50.620		1	144	IPR014729	H	
		Hamap	MF_00151	Phosphopantetheine adenylyltransferase [coaD].	5	145	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342:SF1		5	140	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342		5	140			
		Pfam	PF01467	Cytidylyltransferase-like	8	139	IPR004821	D	GO:0003824; GO:0009058
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	118	140	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	5	23	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	91	107	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	23	44	IPR001980	F	GO:0004595; GO:0015937
		SUPERFAMILY	SSF52374		5	141			

Mycoplasma species	Amino acid region	Database	Database ID	Database signature description	ID region start	ID region end	InterPro ID	ID type	Gene Ontology (GO) term
M. alligatoris	145	TIGRFAM	TIGR01510	coaD_prev_kdtB: pantetheine-phosphate adenylyltransferase	7	139	IPR001980	F	GO:0004595; GO:0015937
		TIGRFAM	TIGR00125	cyt_tran_rel: cytidyltransferase-like domain	7	63	IPR004821	D	GO:0003824; GO:0009058
M. alvi	151	Gene3D	G3DSA:3.40.50.620		1	149	IPR014729	H	
		Hamap	MF_00151	Phosphopantetheine adenylyltransferase [coaD].	4	151	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342		4	140			
		PANTHER	PTHR21342:SF1		4	140	IPR001980	F	GO:0004595; GO:0015937
		Pfam	PF01467	Cytidylyltransferase-like	7	137	IPR004821	D	GO:0003824; GO:0009058
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	22	43	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	116	138	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	89	105	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	4	22	IPR001980	F	GO:0004595; GO:0015937
		SUPERFAMILY	SSF52374		4	143			
		TIGRFAM	TIGR00125	cyt_tran_rel: cytidyltransferase-like domain	5	39	IPR004821	D	GO:0003824; GO:0009058
		TIGRFAM	TIGR01510	coaD_prev_kdtB: pantetheine-phosphate adenylyltransferase	5	140	IPR001980	F	GO:0004595; GO:0015937
M. anatis	143	Gene3D	G3DSA:3.40.50.620		1	140	IPR014729	H	
		Hamap	MF_00151	Phosphopantetheine adenylyltransferase [coaD].	4	143	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342:SF1		4	135	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342		4	135			
		Pfam	PF01467	Cytidylyltransferase-like	7	135	IPR004821	D	GO:0003824; GO:0009058
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	115	137	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	22	43	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	88	104	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	51	75	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	4	22	IPR001980	F	GO:0004595; GO:0015937
		SUPERFAMILY	SSF52374		4	138			

Mycoplasma species	Amino acid region	Database	Database ID	Database signature description	ID region start	ID region end	InterPro ID	ID type	Gene Ontology (GO) term
M. anatis	143	TIGRFAM	TIGR01510	coaD_prev_kdtB: pantetheine-phosphate adenylyltransferase	6	136	IPR001980	F	GO:0004595; GO:0015937
		TIGRFAM	TIGR00125	cyt_tran_rel: cytidyltransferase-like domain	6	46	IPR004821	D	GO:0003824; GO:0009058
M. arginini	143	Gene3D	G3DSA:3.40.50.620		1	142	IPR014729	H	
		Hamap	MF_00151	Phosphopantetheine adenylyltransferase [coaD].	2	143	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342:SF1		1	138	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342		1	138			
		Pfam	PF01467	Cytidylyltransferase-like	5	134	IPR004821	D	GO:0003824; GO:0009058
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	86	102	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	2	20	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	113	135	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	20	41	IPR001980	F	GO:0004595; GO:0015937
		SUPERFAMILY	SSF52374		2	140			
		TIGRFAM	TIGR01510	coaD_prev_kdtB: pantetheine-phosphate adenylyltransferase	4	134	IPR001980	F	GO:0004595; GO:0015937
		TIGRFAM	TIGR00125	cyt_tran_rel: cytidyltransferase-like domain	4	55	IPR004821	D	GO:0003824; GO:0009058
M. bovigenitalium	140	Gene3D	G3DSA:3.40.50.620		1	139	IPR014729	H	
		PANTHER	PTHR21342:SF1		1	136	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342		1	136			
		Pfam	PF01467	Cytidylyltransferase-like	5	134	IPR004821	D	GO:0003824; GO:0009058
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	113	135	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	2	20	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	86	102	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	49	73	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	20	41	IPR001980	F	GO:0004595; GO:0015937
		SUPERFAMILY	SSF52374		2	136			
		TIGRFAM	TIGR01510	coaD_prev_kdtB: pantetheine-phosphate adenylyltransferase	4	134	IPR001980	F	GO:0004595; GO:0015937

Mycoplasma species	Amino acid region	Database	Database ID	Database signature description	ID region start	ID region end	InterPro ID	ID type	Gene Ontology (GO) term
M. bovigenitalium	140	TIGRFAM	TIGR00125	cyt_tran_rel: cytidyltransferase-like domain	4	59	IPR004821	D	GO:0003824; GO:0009058
M. bovis	140	Gene3D	G3DSA:3.40.50.620		1	140	IPR014729	H	
		Hamap	MF_00151	Phosphopantetheine adenylyltransferase [coaD].	2	140	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342		1	139			
		PANTHER	PTHR21342:SF1		1	139	IPR001980	F	GO:0004595; GO:0015937
		Pfam	PF01467	Cytidylyltransferase-like	5	134	IPR004821	D	GO:0003824; GO:0009058
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	49	73	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	2	20	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	20	41	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	113	135	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	86	102	IPR001980	F	GO:0004595; GO:0015937
		SUPERFAMILY	SSF52374		2	139			
		TIGRFAM	TIGR00125	cyt_tran_rel: cytidyltransferase-like domain	4	55	IPR004821	D	GO:0003824; GO:0009058
		TIGRFAM	TIGR01510	coaD_prev_kdtB: pantetheine-phosphate adenylyltransferase	4	134	IPR001980	F	GO:0004595; GO:0015937
M. buteonis	148	Coils	Coil		70	90			
		Gene3D	G3DSA:3.40.50.620		3	147	IPR014729	H	
		Hamap	MF_00151	Phosphopantetheine adenylyltransferase [coaD].	6	148	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342:SF1		6	140	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342		6	140			
		Pfam	PF01467	Cytidylyltransferase-like	9	140	IPR004821	D	GO:0003824; GO:0009058
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	24	45	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	6	24	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	92	108	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	119	141	IPR001980	F	GO:0004595; GO:0015937
		SUPERFAMILY	SSF52374		6	140			

Mycoplasma species	Amino acid region	Database	Database ID	Database signature description	ID region start	ID region end	InterPro ID	ID type	Gene Ontology (GO) term
M. buteonis	148	TIGRFAM	TIGR01510	coaD_prev_kdtB: pantetheine-phosphate adenylyltransferase	7	144	IPR001980	F	GO:0004595; GO:0015937
		TIGRFAM	TIGR00125	cyt_tran_rel: cytidyltransferase-like domain	7	62	IPR004821	D	GO:0003824; GO:0009058
M. californicum	143	Gene3D	G3DSA:3.40.50.620		1	142	IPR014729	H	
		PANTHER	PTHR21342:SF1		1	138	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342		1	138			
		Pfam	PF01467	Cytidylyltransferase-like	5	134	IPR004821	D	GO:0003824; GO:0009058
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	2	20	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	86	102	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	49	73	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	113	135	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	20	41	IPR001980	F	GO:0004595; GO:0015937
		SUPERFAMILY	SSF52374		2	140			
		TIGRFAM	TIGR00125	cyt_tran_rel: cytidyltransferase-like domain	4	57	IPR004821	D	GO:0003824; GO:0009058
		TIGRFAM	TIGR01510	coaD_prev_kdtB: pantetheine-phosphate adenylyltransferase	4	134	IPR001980	F	GO:0004595; GO:0015937
M. canis	143	Coils	Coil		46	66			
		Gene3D	G3DSA:3.40.50.620		1	142	IPR014729	H	
		PANTHER	PTHR21342		5	139			
		PANTHER	PTHR21342:SF1		5	139	IPR001980	F	GO:0004595; GO:0015937
		Pfam	PF01467	Cytidylyltransferase-like	8	137	IPR004821	D	GO:0003824; GO:0009058
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	89	105	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	116	138	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	5	23	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	23	44	IPR001980	F	GO:0004595; GO:0015937
		SUPERFAMILY	SSF52374		5	138			
		TIGRFAM	TIGR01510	coaD_prev_kdtB: pantetheine-phosphate adenylyltransferase	7	139	IPR001980	F	GO:0004595; GO:0015937

Mycoplasma species	Amino acid region	Database	Database ID	Database signature description	ID region start	ID region end	InterPro ID	ID type	Gene Ontology (GO) term
M. canis	143	TIGRFAM	TIGR00125	cyt_tran_rel: cytidyltransferase-like domain	7	47	IPR004821	D	GO:0003824; GO:0009058
M. capricolum	140	Gene3D	G3DSA:3.40.50.620		1	137	IPR014729	H	
		Hamap	MF_00151	Phosphopantetheine adenylyltransferase [coaD].	2	140	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342:SF1		1	138	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342		1	138			
		Pfam	PF01467	Cytidylyltransferase-like	5	134	IPR004821	D	GO:0003824; GO:0009058
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	113	135	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	2	20	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	49	73	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	86	102	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	20	41	IPR001980	F	GO:0004595; GO:0015937
		SUPERFAMILY	SSF52374		2	138			
		TIGRFAM	TIGR01510	coaD_prev_kdtB: pantetheine-phosphate adenylyltransferase	3	139	IPR001980	F	GO:0004595; GO:0015937
		TIGRFAM	TIGR00125	cyt_tran_rel: cytidyltransferase-like domain	3	61	IPR004821	D	GO:0003824; GO:0009058
M. collis	147	Gene3D	G3DSA:3.40.50.620		2	143	IPR014729	H	
		Hamap	MF_00151	Phosphopantetheine adenylyltransferase [coaD].	8	147	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342		7	139			
		PANTHER	PTHR21342:SF1		7	139	IPR001980	F	GO:0004595; GO:0015937
		Pfam	PF01467	Cytidylyltransferase-like	11	136	IPR004821	D	GO:0003824; GO:0009058
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	55	79	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	8	26	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	26	47	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	92	108	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	119	141	IPR001980	F	GO:0004595; GO:0015937
		SUPERFAMILY	SSF52374		8	141			

Mycoplasma species	Amino acid region	Database	Database ID	Database signature description	ID region start	ID region end	InterPro ID	ID type	Gene Ontology (GO) term
M. collis	147	TIGRFAM	TIGR01510	coaD_prev_kdtB: pantetheine-phosphate adenylyltransferase	10	139	IPR001980	F	GO:0004595; GO:0015937
		TIGRFAM	TIGR00125	cyt_tran_rel: cytidyltransferase-like domain	10	67	IPR004821	D	GO:0003824; GO:0009058
M. columbinum	142	Gene3D	G3DSA:3.40.50.620		1	141	IPR014729	H	
		Hamap	MF_00151	Phosphopantetheine adenylyltransferase [coaD].	3	142	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342		1	138			
		PANTHER	PTHR21342:SF1		1	138	IPR001980	F	GO:0004595; GO:0015937
		Pfam	PF01467	Cytidylyltransferase-like	6	136	IPR004821	D	GO:0003824; GO:0009058
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	115	137	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	21	42	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	3	21	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	88	104	IPR001980	F	GO:0004595; GO:0015937
		SUPERFAMILY	SSF52374		3	141			
		TIGRFAM	TIGR00125	cyt_tran_rel: cytidyltransferase-like domain	4	57	IPR004821	D	GO:0003824; GO:0009058
		TIGRFAM	TIGR01510	coaD_prev_kdtB: pantetheine-phosphate adenylyltransferase	4	136	IPR001980	F	GO:0004595; GO:0015937
M. columborale	142	Gene3D	G3DSA:3.40.50.620		1	139	IPR014729	H	
		Hamap	MF_00151	Phosphopantetheine adenylyltransferase [coaD].	4	141	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342		4	136			
		PANTHER	PTHR21342:SF1		4	136	IPR001980	F	GO:0004595; GO:0015937
		Pfam	PF01467	Cytidylyltransferase-like	7	136	IPR004821	D	GO:0003824; GO:0009058
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	115	137	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	4	22	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	88	104	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	51	75	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	22	43	IPR001980	F	GO:0004595; GO:0015937
		SUPERFAMILY	SSF52374		5	136			

Mycoplasma species	Amino acid region	Database	Database ID	Database signature description	ID region start	ID region end	InterPro ID	ID type	Gene Ontology (GO) term
M. columborale	142	TIGRFAM	TIGR00125	cyt_tran_rel: cytidyltransferase-like domain	5	46	IPR004821	D	GO:0003824; GO:0009058
		TIGRFAM	TIGR01510	coaD_prev_kdtB: pantetheine-phosphate adenylyltransferase	5	136	IPR001980	F	GO:0004595; GO:0015937
M. conjunctivae	148	Gene3D	G3DSA:3.40.50.620		2	145	IPR014729	H	
		Hamap	MF_00151	Phosphopantetheine adenylyltransferase [coaD].	11	147	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342:SF1		11	142	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342		11	142			
		Pfam	PF01467	Cytidylyltransferase-like	14	142	IPR004821	D	GO:0003824; GO:0009058
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	95	111	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	122	144	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	29	50	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	11	29	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	58	82	IPR001980	F	GO:0004595; GO:0015937
		SUPERFAMILY	SSF52374		11	144			
		TIGRFAM	TIGR00125	cyt_tran_rel: cytidyltransferase-like domain	13	64	IPR004821	D	GO:0003824; GO:0009058
		TIGRFAM	TIGR01510	coaD_prev_kdtB: pantetheine-phosphate adenylyltransferase	13	142	IPR001980	F	GO:0004595; GO:0015937
M. cricetuli	142	Gene3D	G3DSA:3.40.50.620		1	141	IPR014729	H	
		Hamap	MF_00151	Phosphopantetheine adenylyltransferase [coaD].	4	142	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342:SF1		4	138	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342		4	138			
		Pfam	PF01467	Cytidylyltransferase-like	7	136	IPR004821	D	GO:0003824; GO:0009058
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	88	104	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	115	137	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	51	75	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	22	43	IPR001980	F	GO:0004595; GO:0015937

Mycoplasma species	Amino acid region	Database	Database ID	Database signature description	ID region start	ID region end	InterPro ID	ID type	Gene Ontology (GO) term
M. cricetuli	142	PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	4	22	IPR001980	F	GO:0004595; GO:0015937
		SUPERFAMILY	SSF52374		4	141			
		TIGRFAM	TIGR00125	cyt_tran_rel: cytidyltransferase-like domain	5	59	IPR004821	D	GO:0003824; GO:0009058
		TIGRFAM	TIGR01510	coaD_prev_kdtB: pantetheine-phosphate adenylyltransferase	5	136	IPR001980	F	GO:0004595; GO:0015937
M. crocodyli	146	Gene3D	G3DSA:3.40.50.620		3	146	IPR014729	H	
		Hamap	MF_00151	Phosphopantetheine adenylyltransferase [coaD].	6	146	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342		6	142			
		PANTHER	PTHR21342:SF1		6	142	IPR001980	F	GO:0004595; GO:0015937
		Pfam	PF01467	Cytidylyltransferase-like	9	139	IPR004821	D	GO:0003824; GO:0009058
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	92	108	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	119	141	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	24	45	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	6	24	IPR001980	F	GO:0004595; GO:0015937
		SUPERFAMILY	SSF52374		6	145			
		TIGRFAM	TIGR01510	coaD_prev_kdtB: pantetheine-phosphate adenylyltransferase	8	140	IPR001980	F	GO:0004595; GO:0015937
		TIGRFAM	TIGR00125	cyt_tran_rel: cytidyltransferase-like domain	8	62	IPR004821	D	GO:0003824; GO:0009058
M. felifaucium	146	Coils	Coil		46	66			
		Gene3D	G3DSA:3.40.50.620		1	142	IPR014729	H	
		Hamap	MF_00151	Phosphopantetheine adenylyltransferase [coaD].	4	146	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342:SF1		4	141	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342		4	141			
		Pfam	PF01467	Cytidylyltransferase-like	7	136	IPR004821	D	GO:0003824; GO:0009058
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	51	75	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	22	43	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	4	22	IPR001980	F	GO:0004595; GO:0015937

Mycoplasma species	Amino acid region	Database	Database ID	Database signature description	ID region start	ID region end	InterPro ID	ID type	Gene Ontology (GO) term
M. felifaucium	146	PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	115	137	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	88	104	IPR001980	F	GO:0004595; GO:0015937
		SUPERFAMILY	SSF52374		4	140			
		TIGRFAM	TIGR00125	cyt_tran_rel: cytidyltransferase-like domain	5	62	IPR004821	D	GO:0003824; GO:0009058
		TIGRFAM	TIGR01510	coaD_prev_kdtB: pantetheine-phosphate adenylyltransferase	5	141	IPR001980	F	GO:0004595; GO:0015937
M. felis	137	Gene3D	G3DSA:3.40.50.620		1	135	IPR014729	H	
		PANTHER	PTHR21342		1	133			
		PANTHER	PTHR21342:SF1		1	133	IPR001980	F	GO:0004595; GO:0015937
		Pfam	PF01467	Cytidylyltransferase-like	6	133	IPR004821	D	GO:0003824; GO:0009058
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	21	42	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	3	21	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	86	102	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	113	135	IPR001980	F	GO:0004595; GO:0015937
		SUPERFAMILY	SSF52374		5	134			
		TIGRFAM	TIGR00125	cyt_tran_rel: cytidyltransferase-like domain	5	49	IPR004821	D	GO:0003824; GO:0009058
		TIGRFAM	TIGR01510	coaD_prev_kdtB: pantetheine-phosphate adenylyltransferase	5	133	IPR001980	F	GO:0004595; GO:0015937
M. fermentans	142	Coils	Coil		45	65			
		Gene3D	G3DSA:3.40.50.620		2	142	IPR014729	H	
		Hamap	MF_00151	Phosphopantetheine adenylyltransferase [coaD].	3	142	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342:SF1		1	141	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342		1	141			
		Pfam	PF01467	Cytidylyltransferase-like	6	135	IPR004821	D	GO:0003824; GO:0009058
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	50	74	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	87	103	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	21	42	IPR001980	F	GO:0004595; GO:0015937

Mycoplasma species	Amino acid region	Database	Database ID	Database signature description	ID region start	ID region end	InterPro ID	ID type	Gene Ontology (GO) term
M. fermentans	142	PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	3	21	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	114	136	IPR001980	F	GO:0004595; GO:0015937
		SUPERFAMILY	SSF52374		3	140			
		TIGRFAM	TIGR00125	cyt_tran_rel: cytidyltransferase-like domain	5	58	IPR004821	D	GO:0003824; GO:0009058
		TIGRFAM	TIGR01510	coaD_prev_kdtB: pantetheine-phosphate adenylyltransferase	5	137	IPR001980	F	GO:0004595; GO:0015937
M. gallinaceum	145	Gene3D	G3DSA:3.40.50.620		1	145	IPR014729	H	
		Hamap	MF_00151	Phosphopantetheine adenylyltransferase [coaD].	5	145	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342:SF1		5	144	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342		5	144			
		Pfam	PF01467	Cytidylyltransferase-like	8	137	IPR004821	D	GO:0003824; GO:0009058
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	52	76	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	5	23	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	116	138	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	23	44	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	89	105	IPR001980	F	GO:0004595; GO:0015937
		SUPERFAMILY	SSF52374		5	144			
		TIGRFAM	TIGR00125	cyt_tran_rel: cytidyltransferase-like domain	7	58	IPR004821	D	GO:0003824; GO:0009058
		TIGRFAM	TIGR01510	coaD_prev_kdtB: pantetheine-phosphate adenylyltransferase	7	145	IPR001980	F	GO:0004595; GO:0015937
M. gallinarum	141	Coils	Coil		45	65			
		Gene3D	G3DSA:3.40.50.620		1	139	IPR014729	H	
		Hamap	MF_00151	Phosphopantetheine adenylyltransferase [coaD].	3	141	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342		1	137			
		PANTHER	PTHR21342:SF1		1	137	IPR001980	F	GO:0004595; GO:0015937
		Pfam	PF01467	Cytidylyltransferase-like	6	133	IPR004821	D	GO:0003824; GO:0009058
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	87	103	IPR001980	F	GO:0004595; GO:0015937

Mycoplasma species	Amino acid region	Database	Database ID	Database signature description	ID region start	ID region end	InterPro ID	ID type	Gene Ontology (GO) term
M. gallinarum	141	PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	50	74	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	21	42	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	3	21	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	114	136	IPR001980	F	GO:0004595; GO:0015937
		SUPERFAMILY	SSF52374		3	137			
		TIGRFAM	TIGR01510	coaD_prev_kdtB: pantetheine-phosphate adenylyltransferase	4	140	IPR001980	F	GO:0004595; GO:0015937
		TIGRFAM	TIGR00125	cyt_tran_rel: cytidyltransferase-like domain	4	60	IPR004821	D	GO:0003824; GO:0009058
M. iners	147	Gene3D	G3DSA:3.40.50.620		1	145	IPR014729	H	
		Hamap	MF_00151	Phosphopantetheine adenylyltransferase [coaD].	3	147	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342:SF1		1	136	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342		1	136			
		Pfam	PF01467	Cytidylyltransferase-like	6	132	IPR004821	D	GO:0003824; GO:0009058
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	3	21	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	114	136	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	50	74	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	87	103	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	21	42	IPR001980	F	GO:0004595; GO:0015937
		SUPERFAMILY	SSF52374		3	138			
		TIGRFAM	TIGR01510	coaD_prev_kdtB: pantetheine-phosphate adenylyltransferase	5	140	IPR001980	F	GO:0004595; GO:0015937
		TIGRFAM	TIGR00125	cyt_tran_rel: cytidyltransferase-like domain	4	46	IPR004821	D	GO:0003824; GO:0009058
M. iowae	144	Gene3D	G3DSA:3.40.50.620		1	144	IPR014729	H	
		PANTHER	PTHR21342		6	136			
		PANTHER	PTHR21342:SF1		6	136	IPR001980	F	GO:0004595; GO:0015937
		Pfam	PF01467	Cytidylyltransferase-like	9	137	IPR004821	D	GO:0003824; GO:0009058
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	6	24	IPR001980	F	GO:0004595; GO:0015937

Mycoplasma species	Amino acid region	Database	Database ID	Database signature description	ID region start	ID region end	InterPro ID	ID type	Gene Ontology (GO) term
M. iowae	144	PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	117	139	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	24	45	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	90	106	IPR001980	F	GO:0004595; GO:0015937
		SUPERFAMILY	SSF52374		8	142			
		TIGRFAM	TIGR00125	cyt_tran_rel: cytidyltransferase-like domain	8	63	IPR004821	D	GO:0003824; GO:0009058
		TIGRFAM	TIGR01510	coaD_prev_kdtB: pantetheine-phosphate adenylyltransferase	8	136	IPR001980	F	GO:0004595; GO:0015937
M. leachii	140	Gene3D	G3DSA:3.40.50.620		1	137	IPR014729	H	
		Hamap	MF_00151	Phosphopantetheine adenylyltransferase [coaD].	2	140	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342:SF1		1	138	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342		1	138			
		Pfam	PF01467	Cytidylyltransferase-like	5	134	IPR004821	D	GO:0003824; GO:0009058
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	86	102	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	113	135	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	49	73	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	20	41	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	2	20	IPR001980	F	GO:0004595; GO:0015937
		SUPERFAMILY	SSF52374		2	138			
		TIGRFAM	TIGR00125	cyt_tran_rel: cytidyltransferase-like domain	4	62	IPR004821	D	GO:0003824; GO:0009058
		TIGRFAM	TIGR01510	coaD_prev_kdtB: pantetheine-phosphate adenylyltransferase	3	139	IPR001980	F	GO:0004595; GO:0015937
M. leonicaptivi	137	Gene3D	G3DSA:3.40.50.620		1	136	IPR014729	H	
		PANTHER	PTHR21342:SF1		4	135	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342		4	135			
		Pfam	PF01467	Cytidylyltransferase-like	8	135	IPR004821	D	GO:0003824; GO:0009058
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	22	43	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	87	103	IPR001980	F	GO:0004595; GO:0015937

Mycoplasma species	Amino acid region	Database	Database ID	Database signature description	ID region start	ID region end	InterPro ID	ID type	Gene Ontology (GO) term
M. leonicaptivi	137	PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	4	22	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	114	136	IPR001980	F	GO:0004595; GO:0015937
		ProDom	PD016147	Ligase Lyase Citrate Pro-3S-Lyase Synthetase Transferase Cytidyltransferase-Related:Citrate:Acetate:SH-Citrate Citc Plasmid	5	67			
		SUPERFAMILY	SSF52374		5	135			
		TIGRFAM	TIGR00125	cyt_tran_rel: cytidyltransferase-like domain	5	46	IPR004821	D	GO:0003824; GO:0009058
		TIGRFAM	TIGR01510	coaD_prev_kdtB: pantetheine-phosphate adenylyltransferase	6	135	IPR001980	F	GO:0004595; GO:0015937
M. lipofaciens	140	Coils	Coil		44	64			
		Gene3D	G3DSA:3.40.50.620		1	140	IPR014729	H	
		Hamap	MF_00151	Phosphopantetheine adenylyltransferase [coaD].	2	140	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342:SF1		1	139	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342		1	139			
		Pfam	PF01467	Cytidylyltransferase-like	5	134	IPR004821	D	GO:0003824; GO:0009058
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	2	20	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	113	135	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	49	73	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	86	102	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	20	41	IPR001980	F	GO:0004595; GO:0015937
		SUPERFAMILY	SSF52374		2	139			
		TIGRFAM	TIGR00125	cyt_tran_rel: cytidyltransferase-like domain	4	60	IPR004821	D	GO:0003824; GO:0009058
		TIGRFAM	TIGR01510	coaD_prev_kdtB: pantetheine-phosphate adenylyltransferase	4	134	IPR001980	F	GO:0004595; GO:0015937
M. mobile	145	Coils	Coil		44	64			
		Gene3D	G3DSA:3.40.50.620		2	141	IPR014729	H	
		Pfam	PF01467	Cytidylyltransferase-like	12	72	IPR004821	D	GO:0003824; GO:0009058
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	9	27	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	27	48	IPR001980	F	GO:0004595; GO:0015937

***Mycoplasma* species**	**Amino acid region**	**Database**	**Database ID**	**Database signature description**	**ID region start**	**ID region end**	**InterPro ID**	**ID type**	**Gene Ontology (GO) term**
M. mobile	145	PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	56	80	IPR001980	F	GO:0004595; GO:0015937
		SUPERFAMILY	SSF52374		9	126			
		TIGRFAM	TIGR00125	cyt_tran_rel: cytidyltransferase-like domain	11	61	IPR004821	D	GO:0003824; GO:0009058
M. molare	143	Gene3D	G3DSA:3.40.50.620		1	138	IPR014729	H	
		Hamap	MF_00151	Phosphopantetheine adenylyltransferase [coaD].	2	143	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342:SF1		1	134	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342		1	134			
		Pfam	PF01467	Cytidylyltransferase-like	5	133	IPR004821	D	GO:0003824; GO:0009058
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	113	135	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	49	73	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	2	20	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	86	102	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	20	41	IPR001980	F	GO:0004595; GO:0015937
		ProDom	PD016147	Ligase Lyase Citrate Pro-3S-Lyase Synthetase Transferase Cytidyltransferase-Related:Citrate:Acetate:SH-Citrate Citc Plasmid	11	75			
		SUPERFAMILY	SSF52374		4	138			
		TIGRFAM	TIGR00125	cyt_tran_rel: cytidyltransferase-like domain	4	56	IPR004821	D	GO:0003824; GO:0009058
		TIGRFAM	TIGR01510	coaD_prev_kdtB: pantetheine-phosphate adenylyltransferase	4	134	IPR001980	F	GO:0004595; GO:0015937
M. mycoides subsp. *capri*	140	Gene3D	G3DSA:3.40.50.620		1	137	IPR014729	H	
		Hamap	MF_00151	Phosphopantetheine adenylyltransferase [coaD].	2	140	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342:SF1		1	138	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342		1	138			
		Pfam	PF01467	Cytidylyltransferase-like	5	134	IPR004821	D	GO:0003824; GO:0009058
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	113	135	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	49	73	IPR001980	F	GO:0004595; GO:0015937

***Mycoplasma* species**	**Amino acid region**	**Database**	**Database ID**	**Database signature description**	**ID region start**	**ID region end**	**InterPro ID**	**ID type**	**Gene Ontology (GO) term**
M. mycoides subsp. *capri*	140	PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	2	20	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	86	102	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	20	41	IPR001980	F	GO:0004595; GO:0015937
		SUPERFAMILY	SSF52374		2	138			
		TIGRFAM	TIGR01510	coaD_prev_kdtB: pantetheine-phosphate adenylyltransferase	3	139	IPR001980	F	GO:0004595; GO:0015937
		TIGRFAM	TIGR00125	cyt_tran_rel: cytidyltransferase-like domain	3	61	IPR004821	D	GO:0003824; GO:0009058
M. mycoides subsp. *mycoides*	140	Gene3D	G3DSA:3.40.50.620		1	137	IPR014729	H	
		Hamap	MF_00151	Phosphopantetheine adenylyltransferase [coaD].	2	140	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342:SF1		1	138	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342		1	138			
		Pfam	PF01467	Cytidylyltransferase-like	5	134	IPR004821	D	GO:0003824; GO:0009058
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	49	73	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	113	135	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	2	20	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	20	41	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	86	102	IPR001980	F	GO:0004595; GO:0015937
		SUPERFAMILY	SSF52374		2	138			
		TIGRFAM	TIGR00125	cyt_tran_rel: cytidyltransferase-like domain	4	62	IPR004821	D	GO:0003824; GO:0009058
		TIGRFAM	TIGR01510	coaD_prev_kdtB: pantetheine-phosphate adenylyltransferase	3	139	IPR001980	F	GO:0004595; GO:0015937
M. opalescens	145	Gene3D	G3DSA:3.40.50.620		1	143	IPR014729	H	
		Hamap	MF_00151	Phosphopantetheine adenylyltransferase [coaD].	2	145	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342:SF1		1	139	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342		1	139			
		Pfam	PF01467	Cytidylyltransferase-like	5	134	IPR004821	D	GO:0003824; GO:0009058
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	2	20	IPR001980	F	GO:0004595; GO:0015937

Mycoplasma species	Amino acid region	Database	Database ID	Database signature description	ID region start	ID region end	InterPro ID	ID type	Gene Ontology (GO) term
M. opalescens	145	PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	87	103	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	20	41	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	114	136	IPR001980	F	GO:0004595; GO:0015937
		SUPERFAMILY	SSF52374		2	140			
		TIGRFAM	TIGR00125	cyt_tran_rel: cytidyltransferase-like domain	4	43	IPR004821	D	GO:0003824; GO:0009058
		TIGRFAM	TIGR01510	coaD_prev_kdtB: pantetheine-phosphate adenylyltransferase	4	135	IPR001980	F	GO:0004595; GO:0015937
M. penetrans	150	Coils	Coil		49	69			
		Gene3D	G3DSA:3.40.50.620		2	149	IPR014729	H	
		Hamap	MF_00151	Phosphopantetheine adenylyltransferase [coaD].	8	149	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342		8	136			
		PANTHER	PTHR21342:SF1		8	136	IPR001980	F	GO:0004595; GO:0015937
		Pfam	PF01467	Cytidylyltransferase-like	11	140	IPR004821	D	GO:0003824; GO:0009058
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	8	26	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	119	141	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	92	108	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	26	47	IPR001980	F	GO:0004595; GO:0015937
		SUPERFAMILY	SSF52374		8	147			
		TIGRFAM	TIGR01510	coaD_prev_kdtB: pantetheine-phosphate adenylyltransferase	10	147	IPR001980	F	GO:0004595; GO:0015937
		TIGRFAM	TIGR00125	cyt_tran_rel: cytidyltransferase-like domain	11	65	IPR004821	D	GO:0003824; GO:0009058
M. pirum	151	Gene3D	G3DSA:3.40.50.620		1	149	IPR014729	H	
		Hamap	MF_00151	Phosphopantetheine adenylyltransferase [coaD].	4	151	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342:SF1		4	141	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342		4	141			
		Pfam	PF01467	Cytidylyltransferase-like	7	137	IPR004821	D	GO:0003824; GO:0009058
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	4	22	IPR001980	F	GO:0004595; GO:0015937

Mycoplasma species	Amino acid region	Database	Database ID	Database signature description	ID region start	ID region end	InterPro ID	ID type	Gene Ontology (GO) term
M. pirum	151	PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	22	43	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	89	105	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	116	138	IPR001980	F	GO:0004595; GO:0015937
		SUPERFAMILY	SSF52374		4	143			
		TIGRFAM	TIGR01510	coaD_prev_kdtB: pantetheine-phosphate adenylyltransferase	5	140	IPR001980	F	GO:0004595; GO:0015937
		TIGRFAM	TIGR00125	cyt_tran_rel: cytidyltransferase-like domain	5	62	IPR004821	D	GO:0003824; GO:0009058
M. primatum	140	Gene3D	G3DSA:3.40.50.620		1	140	IPR014729	H	
		Hamap	MF_00151	Phosphopantetheine adenylyltransferase [coaD].	2	140	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342:SF1		1	139	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342		1	139			
		Pfam	PF01467	Cytidylyltransferase-like	5	134	IPR004821	D	GO:0003824; GO:0009058
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	113	135	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	49	73	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	86	102	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	2	20	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	20	41	IPR001980	F	GO:0004595; GO:0015937
		SUPERFAMILY	SSF52374		2	139			
		TIGRFAM	TIGR00125	cyt_tran_rel: cytidyltransferase-like domain	4	59	IPR004821	D	GO:0003824; GO:0009058
		TIGRFAM	TIGR01510	coaD_prev_kdtB: pantetheine-phosphate adenylyltransferase	4	134	IPR001980	F	GO:0004595; GO:0015937
M. pulmonis	149	Gene3D	G3DSA:3.40.50.620		2	143	IPR014729	H	
		Hamap	MF_00151	Phosphopantetheine adenylyltransferase [coaD].	6	148	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342		6	140			
		PANTHER	PTHR21342:SF1		6	140	IPR001980	F	GO:0004595; GO:0015937
		Pfam	PF01467	Cytidylyltransferase-like	9	140	IPR004821	D	GO:0003824; GO:0009058
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	119	141	IPR001980	F	GO:0004595; GO:0015937

Mycoplasma species	Amino acid region	Database	Database ID	Database signature description	ID region start	ID region end	InterPro ID	ID type	Gene Ontology (GO) term
M. pulmonis	149	PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	24	45	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	92	108	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	6	24	IPR001980	F	GO:0004595; GO:0015937
		ProDom	PD016147	Ligase Lyase Citrate Pro-3S-Lyase Synthetase Transferase Cytidyltransferase-Related:Citrate:Acetate:SH-Citrate Citc Plasmid	4	103			
		SUPERFAMILY	SSF52374		6	140			
		TIGRFAM	TIGR01510	coaD_prev_kdtB: pantetheine-phosphate adenylyltransferase	7	140	IPR001980	F	GO:0004595; GO:0015937
		TIGRFAM	TIGR00125	cyt_tran_rel: cytidyltransferase-like domain	7	64	IPR004821	D	GO:0003824; GO:0009058
M. putrefaciens	141	Gene3D	G3DSA:3.40.50.620		1	137	IPR014729	H	
		Hamap	MF_00151	Phosphopantetheine adenylyltransferase [coaD].	2	141	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342		1	139			
		PANTHER	PTHR21342:SF1		1	139	IPR001980	F	GO:0004595; GO:0015937
		Pfam	PF01467	Cytidylyltransferase-like	5	134	IPR004821	D	GO:0003824; GO:0009058
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	86	102	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	113	135	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	20	41	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	2	20	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	49	73	IPR001980	F	GO:0004595; GO:0015937
		ProDom	PD016147	Ligase Lyase Citrate Pro-3S-Lyase Synthetase Transferase Cytidyltransferase-Related:Citrate:Acetate:SH-Citrate Citc Plasmid	2	75			
		SUPERFAMILY	SSF52374		2	138			
		TIGRFAM	TIGR01510	coaD_prev_kdtB: pantetheine-phosphate adenylyltransferase	3	141	IPR001980	F	GO:0004595; GO:0015937
		TIGRFAM	TIGR00125	cyt_tran_rel: cytidyltransferase-like domain	3	62	IPR004821	D	GO:0003824; GO:0009058
M. simbae	142	Gene3D	G3DSA:3.40.50.620		1	138	IPR014729	H	
		PANTHER	PTHR21342		1	136			
		PANTHER	PTHR21342:SF1		1	136	IPR001980	F	GO:0004595; GO:0015937

Mycoplasma species	Amino acid region	Database	Database ID	Database signature description	ID region start	ID region end	InterPro ID	ID type	Gene Ontology (GO) term
M. simbae	142	Pfam	PF01467	Cytidylyltransferase-like	5	134	IPR004821	D	GO:0003824; GO:0009058
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	20	41	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	2	20	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	113	135	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	49	73	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	86	102	IPR001980	F	GO:0004595; GO:0015937
		SUPERFAMILY	SSF52374		2	135			
		TIGRFAM	TIGR01510	coaD_prev_kdtB: pantetheine-phosphate adenylyltransferase	4	134	IPR001980	F	GO:0004595; GO:0015937
		TIGRFAM	TIGR00125	cyt_tran_rel: cytidyltransferase-like domain	4	56	IPR004821	D	GO:0003824; GO:0009058
M. sturni	144	Gene3D	G3DSA:3.40.50.620		1	144	IPR014729	H	
		Hamap	MF_00151	Phosphopantetheine adenylyltransferase [coaD].	4	143	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342:SF1		4	143	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342		4	143			
		Pfam	PF01467	Cytidylyltransferase-like	7	136	IPR004821	D	GO:0003824; GO:0009058
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	4	22	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	22	43	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	115	137	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	51	75	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	88	104	IPR001980	F	GO:0004595; GO:0015937
		SUPERFAMILY	SSF52374		4	142			
		TIGRFAM	TIGR01510	coaD_prev_kdtB: pantetheine-phosphate adenylyltransferase	5	144	IPR001980	F	GO:0004595; GO:0015937
		TIGRFAM	TIGR00125	cyt_tran_rel: cytidyltransferase-like domain	5	62	IPR004821	D	GO:0003824; GO:0009058
M. synoviae	148	Gene3D	G3DSA:3.40.50.620		1	147	IPR014729	H	
		Hamap	MF_00151	Phosphopantetheine adenylyltransferase [coaD].	7	148	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342		7	143			

Mycoplasma species	Amino acid region	Database	Database ID	Database signature description	ID region start	ID region end	InterPro ID	ID type	Gene Ontology (GO) term
M. synoviae	148	PANTHER	PTHR21342:SF1		7	143	IPR001980	F	GO:0004595; GO:0015937
		Pfam	PF01467	Cytidylyltransferase-like	10	139	IPR004821	D	GO:0003824; GO:0009058
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	7	25	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	118	140	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	25	46	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	91	107	IPR001980	F	GO:0004595; GO:0015937
		SUPERFAMILY	SSF52374		7	145			
		TIGRFAM	TIGR00125	cyt_tran_rel: cytidyltransferase-like domain	9	60	IPR004821	D	GO:0003824; GO:0009058
		TIGRFAM	TIGR01510	coaD_prev_kdtB: pantetheine-phosphate adenylyltransferase	9	139	IPR001980	F	GO:0004595; GO:0015937
M. testudinis	151	Gene3D	G3DSA:3.40.50.620		1	149	IPR014729	H	
		Hamap	MF_00151	Phosphopantetheine adenylyltransferase [coaD].	3	150	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342:SF1		1	142	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342		1	142			
		Pfam	PF01467	Cytidylyltransferase-like	6	136	IPR004821	D	GO:0003824; GO:0009058
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	3	21	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	21	42	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	88	104	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	115	137	IPR001980	F	GO:0004595; GO:0015937
		SUPERFAMILY	SSF52374		3	142			
		TIGRFAM	TIGR00125	cyt_tran_rel: cytidyltransferase-like domain	4	61	IPR004821	D	GO:0003824; GO:0009058
		TIGRFAM	TIGR01510	coaD_prev_kdtB: pantetheine-phosphate adenylyltransferase	4	140	IPR001980	F	GO:0004595; GO:0015937
M. yeatsii	140	Gene3D	G3DSA:3.40.50.620		1	138	IPR014729	H	
		Hamap	MF_00151	Phosphopantetheine adenylyltransferase [coaD].	2	140	IPR001980	F	GO:0004595; GO:0015937
		PANTHER	PTHR21342		1	138			
		PANTHER	PTHR21342:SF1		1	138	IPR001980	F	GO:0004595; GO:0015937

Mycoplasma species	Amino acid region	Database	Database ID	Database signature description	ID region start	ID region end	InterPro ID	ID type	Gene Ontology (GO) term
M. yeatsii	140	Pfam	PF01467	Cytidylyltransferase-like	5	134	IPR004821	D	GO:0003824; GO:0009058
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	86	102	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	20	41	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	2	20	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	49	73	IPR001980	F	GO:0004595; GO:0015937
		PRINTS	PR01020	Lipopolysaccharide core biosynthesis protein signature	113	135	IPR001980	F	GO:0004595; GO:0015937
		SUPERFAMILY	SSF52374		2	138			
		TIGRFAM	TIGR00125	cyt_tran_rel: cytidyltransferase-like domain	3	61	IPR004821	D	GO:0003824; GO:0009058
		TIGRFAM	TIGR01510	coaD_prev_kdtB: pantetheine-phosphate adenylyltransferase	3	139	IPR001980	F	GO:0004595; GO:0015937

[a]ID type abbreviations – H, Homologous superfamily; F, Family; D, Domain
[b]InterPro ID – IPR014729: Rossmann-like alpha/beta/alpha sandwich fold; IPR001980: Phosphopantetheine adenylyltransferase; IPR004821: Cytidyltransferase-like domain
[c]GO term (Biological Process) – GO:0009058: Biosynthetic process; GO:0015937: Coenzyme A biosynthetic process
GO term (Molecular Function) – GO:0003824: Catalytic activity; GO:0004595: Pantetheine-phosphate adenylyltransferase activity

Supplementary Table 1.10 The InterPro search results of the DPCK amino acid sequences of the currently annotated CoA biosynthetic pathway enzyme-encoding genes and newly identified homologues

***Mycoplasma* species**	**Amino acid region**	**Database**	**Database ID**	**Database signature description**	**ID region start**	**ID region end**	**InterPro ID**	**ID type**	**Gene Ontology (GO) term**
M. sp. Ms02	192	Gene3D	G3DSA:3.40.50.300		1	144			
		Pfam	PF01121	Dephospho-CoA kinase	1	127	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		SUPERFAMILY	SSF52540		2	106	IPR027417	H	
M. agalactiae	190	CDD	cd02022	DPCK	1	143	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		Gene3D	G3DSA:3.40.50.300		1	163			
		PANTHER	PTHR10695		2	136			
		PANTHER	PTHR10695:SF26		2	136			
		Pfam	PF01121	Dephospho-CoA kinase	1	140	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		SUPERFAMILY	SSF52540		2	137	IPR027417	H	
		TIGRFAM	TIGR00152	TIGR00152: dephospho-CoA kinase	1	141	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
M. alligatoris	185	Gene3D	G3DSA:3.40.50.300		1	183			
		SUPERFAMILY	SSF52540		2	139	IPR027417	H	
M. alvi	196	CDD	cd02022	DPCK	2	151	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		Gene3D	G3DSA:3.40.50.300		1	196			
		PANTHER	PTHR10695:SF35		2	186			
		PANTHER	PTHR10695		2	186			
		Pfam	PF01121	Dephospho-CoA kinase	1	107	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		ProSiteProfiles	PS51219	Dephospho-CoA kinase (DPCK) domain profile.	2	196	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		SUPERFAMILY	SSF52540		1	190	IPR027417	H	
		TIGRFAM	TIGR00152	TIGR00152: dephospho-CoA kinase	2	185	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
M. anatis	190	Gene3D	G3DSA:3.40.50.300		1	187			
		Pfam	PF01121	Dephospho-CoA kinase	1	141	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		SUPERFAMILY	SSF52540		2	137	IPR027417	H	

***Mycoplasma* species**	**Amino acid region**	**Database**	**Database ID**	**Database signature description**	**ID region start**	**ID region end**	**InterPro ID**	**ID type**	**Gene Ontology (GO) term**
M. arginini	168	Gene3D	G3DSA:3.40.50.300		1	167			
		Pfam	PF01121	Dephospho-CoA kinase	2	140	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		SUPERFAMILY	SSF52540		2	124	IPR027417	H	
M. bovigenitalium	190	Gene3D	G3DSA:3.40.50.300		1	190			
		Pfam	PF01121	Dephospho-CoA kinase	2	137	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		SUPERFAMILY	SSF52540		2	181	IPR027417	H	
M. bovis	190	CDD	cd02022	DPCK	1	146	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		Gene3D	G3DSA:3.40.50.300		1	159			
		Pfam	PF01121	Dephospho-CoA kinase	1	144	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		SUPERFAMILY	SSF52540		2	137	IPR027417	H	
M. bovoculi (HAD-DPCK)	444	Coils	Coil		398	418			
		Gene3D	G3DSA:3.40.50.1000		1	261	IPR023214	H	
		Gene3D	G3DSA:3.40.50.300		262	435			
		PANTHER	PTHR10000		1	263			
		PANTHER	PTHR10000		312	406			
		PANTHER	PTHR10000:SF47		1	263			
		PANTHER	PTHR10000:SF47		312	406			
		Pfam	PF08282	haloacid dehalogenase-like hydrolase	7	254			
		Pfam	PF01121	Dephospho-CoA kinase	266	399	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		SUPERFAMILY	SSF52540		264	420	IPR027417	H	
		SUPERFAMILY	SSF56784		1	262	IPR036412	H	
		TIGRFAM	TIGR01484	HAD-SF-IIB: HAD hydrolase, family IIB	7	218	IPR006379	F	
M. buteonis	189	Gene3D	G3DSA:3.40.50.300		1	185			
		SUPERFAMILY	SSF52540		2	179	IPR027417	H	
M. californicum	190	Gene3D	G3DSA:3.40.50.300		15	188			
		Gene3D	G3DSA:3.40.50.300		1	14			
		Pfam	PF01121	Dephospho-CoA kinase	1	146	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		SUPERFAMILY	SSF52540		2	173	IPR027417	H	
M. canis	189	Gene3D	G3DSA:3.40.50.300		1	178			
		SUPERFAMILY	SSF52540		2	141	IPR027417	H	

***Mycoplasma* species**	**Amino acid region**	**Database**	**Database ID**	**Database signature description**	**ID region start**	**ID region end**	**InterPro ID**	**ID type**	**Gene Ontology (GO) term**
M. capricolum subsp. *capricolum*	188	CDD	cd02022	DPCK	6	179	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		Gene3D	G3DSA:3.40.50.300		2	183			
		Hamap	MF_00376	Dephospho-CoA kinase [coaE].	4	188	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		PANTHER	PTHR10695		5	174			
		PANTHER	PTHR10695:SF26		5	174			
		Pfam	PF01121	Dephospho-CoA kinase	5	174	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		ProSiteProfiles	PS51219	Dephospho-CoA kinase (DPCK) domain profile.	6	188	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		SUPERFAMILY	SSF52540		4	174	IPR027417	H	
		TIGRFAM	TIGR00152	TIGR00152: dephospho-CoA kinase	6	178	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
M. capricolum subsp. *capripneumoniae*	184	CDD	cd02022	DPCK	2	175	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		Gene3D	G3DSA:3.40.50.300		1	179			
		PANTHER	PTHR10695:SF26		2	170			
		PANTHER	PTHR10695		2	170			
		Pfam	PF01121	Dephospho-CoA kinase	1	170	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		ProSiteProfiles	PS51219	Dephospho-CoA kinase (DPCK) domain profile.	2	184	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		SUPERFAMILY	SSF52540		1	170	IPR027417	H	
		TIGRFAM	TIGR00152	TIGR00152: dephospho-CoA kinase	2	174	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
M. collis	182	CDD	cd02022	DPCK	1	148	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		Gene3D	G3DSA:3.40.50.300		1	155			
		PANTHER	PTHR10695		2	131			
		PANTHER	PTHR10695:SF26		2	131			
		Pfam	PF01121	Dephospho-CoA kinase	1	100	IPR001977	F	GO:0004140; GO:0005524; GO:0015937

***Mycoplasma* species**	**Amino acid region**	**Database**	**Database ID**	**Database signature description**	**ID region start**	**ID region end**	**InterPro ID**	**ID type**	**Gene Ontology (GO) term**
M. collis	182	ProSiteProfiles	PS51219	Dephospho-CoA kinase (DPCK) domain profile.	1	92	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		SUPERFAMILY	SSF52540		2	136	IPR027417	H	
		TIGRFAM	TIGR00152	TIGR00152: dephospho-CoA kinase	1	150	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
M. columbinum	190	CDD	cd02022	DPCK	1	137	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		Gene3D	G3DSA:3.40.50.300		1	183			
		Pfam	PF01121	Dephospho-CoA kinase	2	138	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		SUPERFAMILY	SSF52540		2	137	IPR027417	H	
		TIGRFAM	TIGR00152	TIGR00152: dephospho-CoA kinase	1	144	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
M. columborale	185	Gene3D	G3DSA:3.40.50.300		1	181			
		Pfam	PF01121	Dephospho-CoA kinase	2	135	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		SUPERFAMILY	SSF52540		1	105	IPR027417	H	
M. conjunctivae (HAD-DPCK)	445	Gene3D	G3DSA:3.40.50.1000		1	266	IPR023214	H	
		Gene3D	G3DSA:3.40.50.300		267	439			
		PANTHER	PTHR10000:SF47		341	412			
		PANTHER	PTHR10000:SF47		4	262			
		PANTHER	PTHR10000		341	412			
		PANTHER	PTHR10000		4	262			
		Pfam	PF08282	haloacid dehalogenase-like hydrolase	7	257			
		Pfam	PF01121	Dephospho-CoA kinase	267	406	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		SUPERFAMILY	SSF52540		266	401	IPR027417	H	
		SUPERFAMILY	SSF56784		1	261	IPR036412	H	
		TIGRFAM	TIGR01484	HAD-SF-IIB: HAD hydrolase, family IIB	6	220	IPR006379	F	
M. cricetuli	191	Gene3D	G3DSA:3.40.50.300		1	186			
		SUPERFAMILY	SSF52540		2	149	IPR027417	H	
M. crocodyli	190	Gene3D	G3DSA:3.40.50.300		1	190			
		Pfam	PF01121	Dephospho-CoA kinase	1	157	IPR001977	F	GO:0004140; GO:0005524; GO:0015937

***Mycoplasma* species**	**Amino acid region**	**Database**	**Database ID**	**Database signature description**	**ID region start**	**ID region end**	**InterPro ID**	**ID type**	**Gene Ontology (GO) term**
M. crocodyli	190	ProSiteProfiles	PS51219	Dephospho-CoA kinase (DPCK) domain profile.	1	190	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		SUPERFAMILY	SSF52540		2	149	IPR027417	H	
M. dispar (HAD-DPCK)	447	Gene3D	G3DSA:3.40.50.1000		2	267	IPR023214	H	
		Gene3D	G3DSA:3.40.50.300		268	431			
		PANTHER	PTHR10000		357	403			
		PANTHER	PTHR10000		1	267			
		PANTHER	PTHR10000:SF47		357	403			
		PANTHER	PTHR10000:SF47		1	267			
		Pfam	PF08282	haloacid dehalogenase-like hydrolase	7	258			
		Pfam	PF01121	Dephospho-CoA kinase	269	399	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		SUPERFAMILY	SSF52540		266	401	IPR027417	H	
		SUPERFAMILY	SSF56784		3	263	IPR036412	H	
		TIGRFAM	TIGR01484	HAD-SF-IIB: HAD hydrolase, family IIB	7	230	IPR006379	F	
M. felifaucium	190	Gene3D	G3DSA:3.40.50.300		1	187			
		Pfam	PF01121	Dephospho-CoA kinase	1	137	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		SUPERFAMILY	SSF52540		2	180	IPR027417	H	
M. felis	184	Gene3D	G3DSA:3.40.50.300		1	181			
		Pfam	PF01121	Dephospho-CoA kinase	4	118	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		SUPERFAMILY	SSF52540		1	133	IPR027417	H	
M. fermentans	190	CDD	cd02022	DPCK	1	146	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		Gene3D	G3DSA:3.40.50.300		1	181			
		PANTHER	PTHR10695		2	143			
		PANTHER	PTHR10695:SF26		2	143			
		Pfam	PF01121	Dephospho-CoA kinase	1	146	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		ProSiteProfiles	PS51219	Dephospho-CoA kinase (DPCK) domain profile.	1	190	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		SUPERFAMILY	SSF52540		2	142	IPR027417	H	
		TIGRFAM	TIGR00152	TIGR00152: dephospho-CoA kinase	1	146	IPR001977	F	GO:0004140; GO:0005524; GO:0015937

***Mycoplasma* species**	**Amino acid region**	**Database**	**Database ID**	**Database signature description**	**ID region start**	**ID region end**	**InterPro ID**	**ID type**	**Gene Ontology (GO) term**
M. flocculare (HAD-DPCK)	447	Gene3D	G3DSA:3.40.50.300		266	439			
		Gene3D	G3DSA:3.40.50.1000		2	251	IPR023214	H	
		PANTHER	PTHR10000		1	264			
		PANTHER	PTHR10000		364	382			
		PANTHER	PTHR10000:SF47		1	264			
		PANTHER	PTHR10000:SF47		364	382			
		Pfam	PF08282	haloacid dehalogenase-like hydrolase	7	257			
		SUPERFAMILY	SSF56784		3	263	IPR036412	H	
		SUPERFAMILY	SSF52540		268	402	IPR027417	H	
		TIGRFAM	TIGR01484	HAD-SF-IIB: HAD hydrolase, family IIB	7	229	IPR006379	F	
M. gallinarum	189	Gene3D	G3DSA:3.40.50.300		1	175			
		Pfam	PF01121	Dephospho-CoA kinase	1	151	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		SUPERFAMILY	SSF52540		2	141	IPR027417	H	
M. gallisepticum	200	Gene3D	G3DSA:3.40.50.300		1	199			
		Hamap	MF_00376	Dephospho-CoA kinase [coaE].	7	197	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		PANTHER	PTHR10695:SF26		8	190			
		PANTHER	PTHR10695		8	190			
		Pfam	PF01121	Dephospho-CoA kinase	9	90	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		ProSiteProfiles	PS51219	Dephospho-CoA kinase (DPCK) domain profile.	9	200	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		SignalP_GRAM_POSITIVE	SignalP-TM		1	12			
		SUPERFAMILY	SSF52540		8	195	IPR027417	H	
		TIGRFAM	TIGR00152	TIGR00152: dephospho-CoA kinase	9	188	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
M. genitalium	198	CDD	cd02022	DPCK	2	169	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		Gene3D	G3DSA:3.40.50.300		1	189			
		Hamap	MF_00376	Dephospho-CoA kinase [coaE].	1	188	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		PANTHER	PTHR10695:SF26		2	132			
		PANTHER	PTHR10695		2	132			

***Mycoplasma* species**	**Amino acid region**	**Database**	**Database ID**	**Database signature description**	**ID region start**	**ID region end**	**InterPro ID**	**ID type**	**Gene Ontology (GO) term**
M. genitalium	198	Pfam	PF01121	Dephospho-CoA kinase	1	155	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		ProSiteProfiles	PS51219	Dephospho-CoA kinase (DPCK) domain profile.	2	90	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		SUPERFAMILY	SSF52540		1	186	IPR027417	H	
		TIGRFAM	TIGR00152	TIGR00152: dephospho-CoA kinase	2	180	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
M. hyopneumoniae (HAD-DPCK)	446	Gene3D	G3DSA:3.40.50.300		269	429			
		Gene3D	G3DSA:3.40.50.1000		1	249	IPR023214	H	
		PANTHER	PTHR10000		1	266			
		PANTHER	PTHR10000		341	354			
		PANTHER	PTHR10000:SF47		1	266			
		PANTHER	PTHR10000:SF47		341	354			
		Pfam	PF08282	haloacid dehalogenase-like hydrolase	7	258			
		Pfam	PF01121	Dephospho-CoA kinase	270	428	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		SUPERFAMILY	SSF56784		2	263	IPR036412	H	
		SUPERFAMILY	SSF52540		269	402	IPR027417	H	
		TIGRFAM	TIGR01484	HAD-SF-IIB: HAD hydrolase, family IIB	7	227	IPR006379	F	
M. hyorhinis	195	Gene3D	G3DSA:3.40.50.300		1	184			
		Pfam	PF01121	Dephospho-CoA kinase	12	154	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		SUPERFAMILY	SSF52540		5	158	IPR027417	H	
M. imitans	196	CDD	cd02022	DPCK	9	182	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		Gene3D	G3DSA:3.40.50.300		1	196			
		PANTHER	PTHR10695:SF26		7	190			
		PANTHER	PTHR10695		7	190			
		Pfam	PF01121	Dephospho-CoA kinase	9	100	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		Phobius	SIGNAL_PEPTIDE_C_REGION	C-terminal region of a signal peptide.	14	18			
		Phobius	SIGNAL_PEPTIDE	Signal peptide region	1	18			
		Phobius	NON_CYTOPLASMIC_DOMAIN	Region of a membrane-bound protein predicted to be outside the membrane, in the extracellular region.	19	196			
		Phobius	SIGNAL_PEPTIDE_N_REGION	N-terminal region of a signal peptide.	1	7			
		Phobius	SIGNAL_PEPTIDE_H_REGION	Hydrophobic region of a signal peptide.	8	13			

Mycoplasma species	Amino acid region	Database	Database ID	Database signature description	ID region start	ID region end	InterPro ID	ID type	Gene Ontology (GO) term
M. imitans	196	ProSiteProfiles	PS51219	Dephospho-CoA kinase (DPCK) domain profile.	9	196	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		SUPERFAMILY	SSF52540		8	194	IPR027417	H	
		TIGRFAM	TIGR00152	TIGR00152: dephospho-CoA kinase	9	189	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
M. iners	190	Gene3D	G3DSA:3.40.50.300		1	187			
		Pfam	PF01121	Dephospho-CoA kinase	2	140	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		SUPERFAMILY	SSF52540		2	137	IPR027417	H	
M. iowae	203	CDD	cd02022	DPCK	11	189	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		Gene3D	G3DSA:3.40.50.300		1	203			
		Hamap	MF_00376	Dephospho-CoA kinase [coaE].	9	200	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		PANTHER	PTHR10695		11	199			
		PANTHER	PTHR10695:SF35		11	199			
		Pfam	PF01121	Dephospho-CoA kinase	11	189	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		ProSiteProfiles	PS51219	Dephospho-CoA kinase (DPCK) domain profile.	11	203	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		SUPERFAMILY	SSF52540		8	200	IPR027417	H	
		TIGRFAM	TIGR00152	TIGR00152: dephospho-CoA kinase	11	195	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
M. leachii	185	CDD	cd02022	DPCK	2	175	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		Gene3D	G3DSA:3.40.50.300		1	181			
		PANTHER	PTHR10695:SF26		2	159			
		PANTHER	PTHR10695		2	159			
		Pfam	PF01121	Dephospho-CoA kinase	1	163	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		ProSiteProfiles	PS51219	Dephospho-CoA kinase (DPCK) domain profile.	2	185	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		SUPERFAMILY	SSF52540		1	159	IPR027417	H	

***Mycoplasma* species**	**Amino acid region**	**Database**	**Database ID**	**Database signature description**	**ID region start**	**ID region end**	**InterPro ID**	**ID type**	**Gene Ontology (GO) term**
M. leachii	185	TIGRFAM	TIGR00152	TIGR00152: dephospho-CoA kinase	2	163	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
M. leonicaptivi	187	Gene3D	G3DSA:3.40.50.300		1	183			
		PANTHER	PTHR10695:SF26		2	120			
		PANTHER	PTHR10695		2	120			
		Pfam	PF01121	Dephospho-CoA kinase	1	120	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		SUPERFAMILY	SSF52540		2	116	IPR027417	H	
M. lipofaciens	190	Gene3D	G3DSA:3.40.50.300		1	190			
		Pfam	PF01121	Dephospho-CoA kinase	2	137	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		SUPERFAMILY	SSF52540		2	124	IPR027417	H	
M. mobile	187	Gene3D	G3DSA:3.40.50.300		1	186			
		Pfam	PF01121	Dephospho-CoA kinase	1	156	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		SUPERFAMILY	SSF52540		2	180	IPR027417	H	
M. molare	182	Gene3D	G3DSA:3.40.50.300		1	177			
		Pfam	PF01121	Dephospho-CoA kinase	1	135	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		ProSiteProfiles	PS51219	Dephospho-CoA kinase (DPCK) domain profile.	1	182	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		SUPERFAMILY	SSF52540		2	137	IPR027417	H	
		TIGRFAM	TIGR00152	TIGR00152: dephospho-CoA kinase	1	153	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
M. mycoides subsp. *capri*	185	CDD	cd02022	DPCK	2	175	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		Gene3D	G3DSA:3.40.50.300		1	179			
		PANTHER	PTHR10695:SF26		2	170			
		PANTHER	PTHR10695		2	170			
		Pfam	PF01121	Dephospho-CoA kinase	1	170	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		ProSiteProfiles	PS51219	Dephospho-CoA kinase (DPCK) domain profile.	2	185	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		SUPERFAMILY	SSF52540		1	170	IPR027417		

***Mycoplasma* species**	**Amino acid region**	**Database**	**Database ID**	**Database signature description**	**ID region start**	**ID region end**	**InterPro ID**	**ID type**	**Gene Ontology (GO) term**
M. mycoides subsp. *capri*	185	TIGRFAM	TIGR00152	TIGR00152: dephospho-CoA kinase	2	174	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
M. mycoides subsp. *mycoides*	189	CDD	cd02022	DPCK	6	179	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		Gene3D	G3DSA:3.40.50.300		2	185			
		Hamap	MF_00376	Dephospho-CoA kinase [coaE].	4	188	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		PANTHER	PTHR10695		5	163			
		PANTHER	PTHR10695:SF26		5	163			
		Pfam	PF01121	Dephospho-CoA kinase	5	167	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		ProSiteProfiles	PS51219	Dephospho-CoA kinase (DPCK) domain profile.	6	189	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		SUPERFAMILY	SSF52540		4	163	IPR027417	H	
		TIGRFAM	TIGR00152	TIGR00152: dephospho-CoA kinase	6	167	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
M. opalescens	187	Gene3D	G3DSA:3.40.50.300		1	180			
		Pfam	PF01121	Dephospho-CoA kinase	1	132	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		SUPERFAMILY	SSF52540		2	137	IPR027417	H	
M. ovipneumoniae (HAD-DPCK)	447	Gene3D	G3DSA:3.40.50.1000		2	267	IPR023214	H	
		Gene3D	G3DSA:3.40.50.300		268	431			
		PANTHER	PTHR10000		2	264			
		Pfam	PF08282	haloacid dehalogenase-like hydrolase	7	259			
		Pfam	PF01121	Dephospho-CoA kinase	269	369	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		SUPERFAMILY	SSF56784		2	263	IPR036412	H	
		SUPERFAMILY	SSF52540		266	410	IPR027417	H	
		TIGRFAM	TIGR01484	HAD-SF-IIB: HAD hydrolase, family IIB	6	229	IPR006379	F	
M. penetrans	206	CDD	cd02022	DPCK	14	143	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		Gene3D	G3DSA:3.40.50.300		3	153			
		Hamap	MF_00376	Dephospho-CoA kinase [coaE].	12	206	IPR001977	F	GO:0004140; GO:0005524; GO:0015937

Mycoplasma species	Amino acid region	Database	Database ID	Database signature description	ID region start	ID region end	InterPro ID	ID type	Gene Ontology (GO) term
M. penetrans	206	PANTHER	PTHR10695:SF35		14	200			
		PANTHER	PTHR10695		14	200			
		Pfam	PF01121	Dephospho-CoA kinase	14	144	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		ProSiteProfiles	PS51219	Dephospho-CoA kinase (DPCK) domain profile.	14	206	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		SUPERFAMILY	SSF52540		13	201	IPR027417	H	
		TIGRFAM	TIGR00152	TIGR00152: dephospho-CoA kinase	14	193	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
M. pirum	194	CDD	cd02022	DPCK	7	179	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		Gene3D	G3DSA:3.40.50.300		1	194			
		PANTHER	PTHR10695:SF26		5	192			
		PANTHER	PTHR10695		5	192			
		Pfam	PF01121	Dephospho-CoA kinase	6	138	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		ProSiteProfiles	PS51219	Dephospho-CoA kinase (DPCK) domain profile.	7	194	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		SUPERFAMILY	SSF52540		6	192	IPR027417	H	
		TIGRFAM	TIGR00152	TIGR00152: dephospho-CoA kinase	7	185	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
M. pneumoniae	200	CDD	cd02022	DPCK	2	156	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		Gene3D	G3DSA:3.40.50.300		1	199			
		Hamap	MF_00376	Dephospho-CoA kinase [coaE].	1	188	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		PANTHER	PTHR10695		2	156			
		PANTHER	PTHR10695:SF26		2	156			
		Pfam	PF01121	Dephospho-CoA kinase	1	156	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		ProSiteProfiles	PS51219	Dephospho-CoA kinase (DPCK) domain profile.	2	200	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		SUPERFAMILY	SSF52540		1	186	IPR027417	H	

Mycoplasma species	Amino acid region	Database	Database ID	Database signature description	ID region start	ID region end	InterPro ID	ID type	Gene Ontology (GO) term
M. pneumoniae	200	TIGRFAM	TIGR00152	TIGR00152: dephospho-CoA kinase	2	182	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
M. primatum	190	Gene3D	G3DSA:3.40.50.300		1	188			
		Pfam	PF01121	Dephospho-CoA kinase	1	137	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		SUPERFAMILY	SSF52540		2	188	IPR027417	H	
M. pulmonis	172	CDD	cd02022	DPCK	1	170	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		Gene3D	G3DSA:3.40.50.300		1	171			
		Pfam	PF01121	Dephospho-CoA kinase	1	143	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		ProSiteProfiles	PS51219	Dephospho-CoA kinase (DPCK) domain profile.	1	172	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		SUPERFAMILY	SSF52540		2	137	IPR027417	H	
		TIGRFAM	TIGR00152	TIGR00152: dephospho-CoA kinase	1	143	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
M. putrefaciens	186	CDD	cd02022	DPCK	2	168	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		Gene3D	G3DSA:3.40.50.300		1	185			
		PANTHER	PTHR10695		2	149			
		PANTHER	PTHR10695:SF26		2	149			
		Pfam	PF01121	Dephospho-CoA kinase	1	158	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		ProSiteProfiles	PS51219	Dephospho-CoA kinase (DPCK) domain profile.	2	186	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		SUPERFAMILY	SSF52540		1	161	IPR027417	H	
		TIGRFAM	TIGR00152	TIGR00152: dephospho-CoA kinase	2	162	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
M. simbae	187	Gene3D	G3DSA:3.40.50.300		1	187			
		Pfam	PF01121	Dephospho-CoA kinase	1	144	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		SUPERFAMILY	SSF52540		2	157	IPR027417	H	

Mycoplasma species	Amino acid region	Database	Database ID	Database signature description	ID region start	ID region end	InterPro ID	ID type	Gene Ontology (GO) term
M. simbae	187	TIGRFAM	TIGR00152	TIGR00152: dephospho-CoA kinase	1	149	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
M. sturni	187	Gene3D	G3DSA:3.40.50.300		1	187			
		SUPERFAMILY	SSF52540		2	117	IPR027417	H	
M. synoviae	168	Gene3D	G3DSA:3.40.50.300		1	161			
		SUPERFAMILY	SSF52540		2	124	IPR027417	H	
M. testudinis	205	CDD	cd02022	DPCK	2	175	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		Gene3D	G3DSA:3.40.50.300		1	198			
		PANTHER	PTHR10695:SF35		2	185			
		PANTHER	PTHR10695		2	185			
		Pfam	PF01121	Dephospho-CoA kinase	1	151	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		ProSiteProfiles	PS51219	Dephospho-CoA kinase (DPCK) domain profile.	2	201	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		SUPERFAMILY	SSF52540		1	188	IPR027417	H	
		TIGRFAM	TIGR00152	TIGR00152: dephospho-CoA kinase	2	173	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
M. yeatsii	187	CDD	cd02022	DPCK	2	169	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		Gene3D	G3DSA:3.40.50.300		1	187			
		PANTHER	PTHR10695		2	115			
		PANTHER	PTHR10695:SF26		2	115			
		Pfam	PF01121	Dephospho-CoA kinase	1	163	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		ProSiteProfiles	PS51219	Dephospho-CoA kinase (DPCK) domain profile.	2	187	IPR001977	F	GO:0004140; GO:0005524; GO:0015937
		SUPERFAMILY	SSF52540		1	163	IPR027417	H	
		TIGRFAM	TIGR00152	TIGR00152: dephospho-CoA kinase	2	167	IPR001977	F	GO:0004140; GO:0005524; GO:0015937

[a]ID type abbreviations – H, Homologous superfamily; F, Family

[b]InterPro ID – IPR027417: P-loop containing nucleoside triphosphate hydrolase; IPR001977: Dephospho-CoA kinase; IPR036412: HAD-like superfamily; IPR023214: HAD superfamily; IPR006379: HAD-superfamily hydrolase, subfamily IIB

[c]GO term (Biological Process) – GO:0015937: Coenzyme A biosynthetic process

GO term (Molecular Function) – GO:0004140: Dephospho-CoA kinase activity; GO:0005524: ATP binding

Supplementary Table 1.11 The MEME-predicted motif consensus sequences of the respective CoA biosynthetic pathway enzyme-encoding gene amino acid sequences

Protein	Motif	Motif size	E-value	Regular expression (consensus sequence)
PanK type III	1	15	7.1e-056	[LI]VID[IV]GN[ST][YN][ILT]K[IF][GA][IL]F
	2	29	1.3e-132	IGND[IL]L[AG][LA][AS]EY[ACV]SX[KE]NKNA[IL][AGI][FI]S[FL]GTA[ST]V
	3	29	2.0e-072	[ND][NK]K[LIF][EK]G[AV][SI]IAPG[IL]GFS[FY]N[NAS]L[IL]SKA[SK][KL][LI][KN][KS]
	4	21	3.0e-059	YGT[ND]T[IQ][DE]AL[EN]SGY[NY][NH]L[RK][NS]GFI
CoaBC	1	40	1.2e-104	[SP][EDQ]H[IV][EY][LA][AV][KR][WDE][AIN]D[LR][IF][IL][VL][YAF][PA]A[TS][YFA]NT[IL][NST]K[FIVY][AS][NR]GI[ACN]DN[FI][IALV][TL][SAT][IL]L[SA]
	2	28	2.8e-094	G[HV][IT]F[IL]GP[DEN][CFKY]G[LMQT]L[YK][CDE][GN][DEY][IEF]GIG[RK][VM][ASV][EK][PV][EDS][DE][IV]
	3	29	1.0e-082	[KP]K[ILV]L[IL][ST][FYI]G[YA][TS]K[VIS][KYP][LI]D[PDS][VI]R[ST][VIL][SMT]VPSSGK[MFS]G
	4	50	1.1e-080	[LV]TVIN[AG][NS]L[PQ]HLN[NS][YK]IP[SA]NVK[VI][IV]N[VI][NS][STY][VI][DN]EY[FKY][EK]AV[FD][KS][EY]I[KV][NK][NS]D[GI][FY]IS[VL][AC]AV
PPAT	1	42	2.4e-1027	[MK]K[IK]AI[YF][PA]GSFDP[FL]HKGH[LI][NS][IV][LI]KKALK[LI]FDK[VL]Y[VIL][VI]VSINPDKN
	2	15	3.1e-070	[IL][ED]ER[YF]XN[VI]K[KN]K[LI]K[ND][FL]
	3	21	6.8e-253	N[VI]E[VI][LI]IN[KE][ND][KE]L[TI]A[NE][IL]AK[EK]LN[VA]
	4	50	2.8e-1049	[FY][LI][IV]R[SG]ARN[ND]XD[FY]NYE[LI]ELAAGNKS[LI]N[PN][ND][LI]ET[IV]L[IF][IM][PS]DYEMIN[YI]SSTL[IEL]RH
DPCK	1	21	2.5e-482	MIA[IV][IVT]GKIG[SV]GK[TS]T[FL]L[NK]KL[EK]K
	2	21	6.6e-274	G[YF][KS][VIT][FL][NY][AC]D[ED]F[VI][KH][KN]LY[EK]KNNF[CG]
	3	16	1.2e-142	[ED]N[GK]V[DS]KKK[IL][RK]K[WL][IL]FE[ND]
	4	32	2.7e-389	EHLKNNK[YF]DFVEIPNLX[ST]KNA[DN]FS[KS][LF]FSK[IV][IL][CN]

Supplementary Table 1.12 The MEME-predicted motif regions in the PanK type III amino acid sequences

***Mycoplasma* species**	***p*-value**	**Motif locations**			
		Motif 1	**Motif 2**	**Motif 3**	**Motif 4**
M. alligatoris	4.92e-61	9-23	106-134	140-168	179-199
M. alvi	1.57e-56	9-23	110-138	144-172	183-203
M. anatis	4.52e-15	8-22	-	146-174	-
M. arginini	3.89e-62	10-24	109-137	143-171	182-202
M. buteonis	2.10e-16	5-19	-	134-162	-
M. columborale	9.16e-54	5-19	112-140	146-174	189-209
M. cricetuli	2.89e-52	5-19	112-140	146-174	189-209
M. crocodyli	5.10e-58	14-28	111-139	145-173	184-204
M. gallinaceum	2.33e-56	6-20	101-129	135-163	174-194
M. iowae	4.98e-57	6-20	106-134	140-168	179-199
M. mobile	9.91e-41	3-17	100-128	134-162	-
M. molare	4.05e-34	2-16	100-128	134-162	-
M. penetrans	1.23e-54	9-23	108-136	142-170	181-201
M. pirum	1.09e-53	9-23	108-136	142-170	181-201
M. pulmonis	1.02e-33	3-17	105-133	139-167	-
M. sturni	2.48e-45	5-19	112-140	146-174	189-209
M. synoviae	1.86e-62	10-24	109-137	143-171	182-202
M. testudinis	8.30e-46	5-19	104-132	138-166	-

Supplementary Table 1.13 The MEME-predicted motif regions in the CoaBC amino acid sequences, including the separate PPCS and PPCDC amino acid sequences of *M. iowae*

***Mycoplasma* species**	***p*-value**	**Motif locations**			
		Motif 1	**Motif 2**	**Motif 3**	**Motif 4**
M. anatis	2.69e-117	61-100	134-161	171-199	212-261
M. arginini	1.95e-121	66-105	138-165	175-203	216-265
M. columborale	6.19e-120	61-100	134-161	171-199	212-261
M. cricetuli	8.35e-115	61-100	134-161	171-199	212-261
M. gallinaceum	1.20e-112	68-107	141-168	178-206	219-268
M. mobile	1.77e-60	62-101	134-161	176-204	-
M. sturni	5.66e-122	61-100	134-161	171-199	212-261
M. synoviae	3.78e-121	66-105	138-165	175-205	216-265
M. testudinis	1.42e-56	84-123	157-184	198-226	-
M. iowae (PPCDC)	5.76e-35	69-108	143-170	-	-
M. iowae (PPCS)	5.34e-14	-	-	1-29	-

Supplementary Table 1.14 The MEME-predicted motif regions in the PPAT amino acid sequences

***Mycoplasma* species**	***p*-value**	**Motif locations**			
		Motif 1	**Motif 2**	**Motif 3**	**Motif 4**
M. agalactiae	8.04e-99	1-42	47-61	63-83	85-134
M. alligatoris	4.76e-92	4-45	50-64	68-88	90-139
M. alvi	4.46e-57	3-44	49-63	-	88-137
M. anatis	8.32e-97	3-44	49-63	65-85	87-136
M. arginini	9.84e-101	1-42	47-61	63-83	85-134
M. bovigenitalium	1.33e-88	1-42	47-61	63-83	85-134
M. bovis	9.01e-101	1-42	47-61	63-83	85-134
M. buteonis	3.59e-77	5-46	51-65	69-89	91-140
M. californicum	6.44e-91	1-42	47-61	63-83	85-134
M. canis	5.39e-74	4-45	49-63	66-86	88-137
M. capricolum	1.01e-89	1-42	47-61	63-83	85-134
M. collis	1.04e-81	7-48	53-67	69-89	91-140
M. columbinum	1.37e-97	2-43	48-62	65-85	87-136
M. columborale	4.20e-89	3-44	49-63	65-85	87-136
M. conjunctivae	2.15e-95	10-51	56-70	72-92	94-143
M. cricetuli	4.57e-95	3-44	49-63	65-85	87-136
M. crocodyli	3.14e-91	5-46	51-65	69-89	91-140
M. felifaucium	7.02e-97	3-44	49-63	65-85	87-136
M. felis	9.55e-85	2-43	47-61	63-83	85-134
M. fermentans	8.34e-94	2-43	48-62	64-84	86-135
M. gallinaceum	6.43e-91	4-45	50-64	66-86	88-137
M. gallinarum	1.52e-99	2-43	48-62	64-84	86-135
M. iners	4.14e-87	2-43	48-62	64-84	86-135
M. iowae	1.23e-60	5-46	50-64	-	89-138
M. leachii	6.70e-90	1-42	47-61	63-83	85-134
M. leonicaptivi	6.91e-72	3-44	48-62	64-84	86-135
M. lipofaciens	2.75e-95	1-42	47-61	63-83	85-134
M. mobile	7.89e-27	8-49	54-68	-	-
M. molare	1.33e-89	1-42	47-61	63-83	85-134
M. mycoides subsp. *capri*	1.37e-89	1-42	47-61	63-83	85-134
M. mycoides subsp. *mycoides*	1.05e-87	1-42	47-61	63-83	85-134
M. opalescens	5.58e-90	1-42	47-61	64-84	86-135

***Mycoplasma* species**	***p*-value**	**Motif locations**			
		Motif 1	**Motif 2**	**Motif 3**	**Motif 4**
M. penetrans	4.73e-53	7-48	52-66	-	91-140
M. pirum	1.88e-59	3-44	49-63	-	88-137
M. primatum	7.48e-105	1-42	47-61	63-83	85-134
M. pulmonis	5.26e-87	5-46	51-65	69-89	91-140
M. putrefaciens	4.92e-74	1-42	47-61	63-83	85-134
M. simbae	2.98e-93	1-42	47-61	63-83	85-134
M. sturni	1.22e-91	3-44	49-63	65-85	87-136
M. synoviae	5.65e-98	6-47	52-66	68-88	90-139
M. testudinis	1.46e-63	2-43	48-62	-	87-136
M. yeatsii	1.25e-75	1-42	47-61	63-83	85-134

Supplementary Table 1.15 The MEME-predicted motif regions in the DPCK amino acid sequences

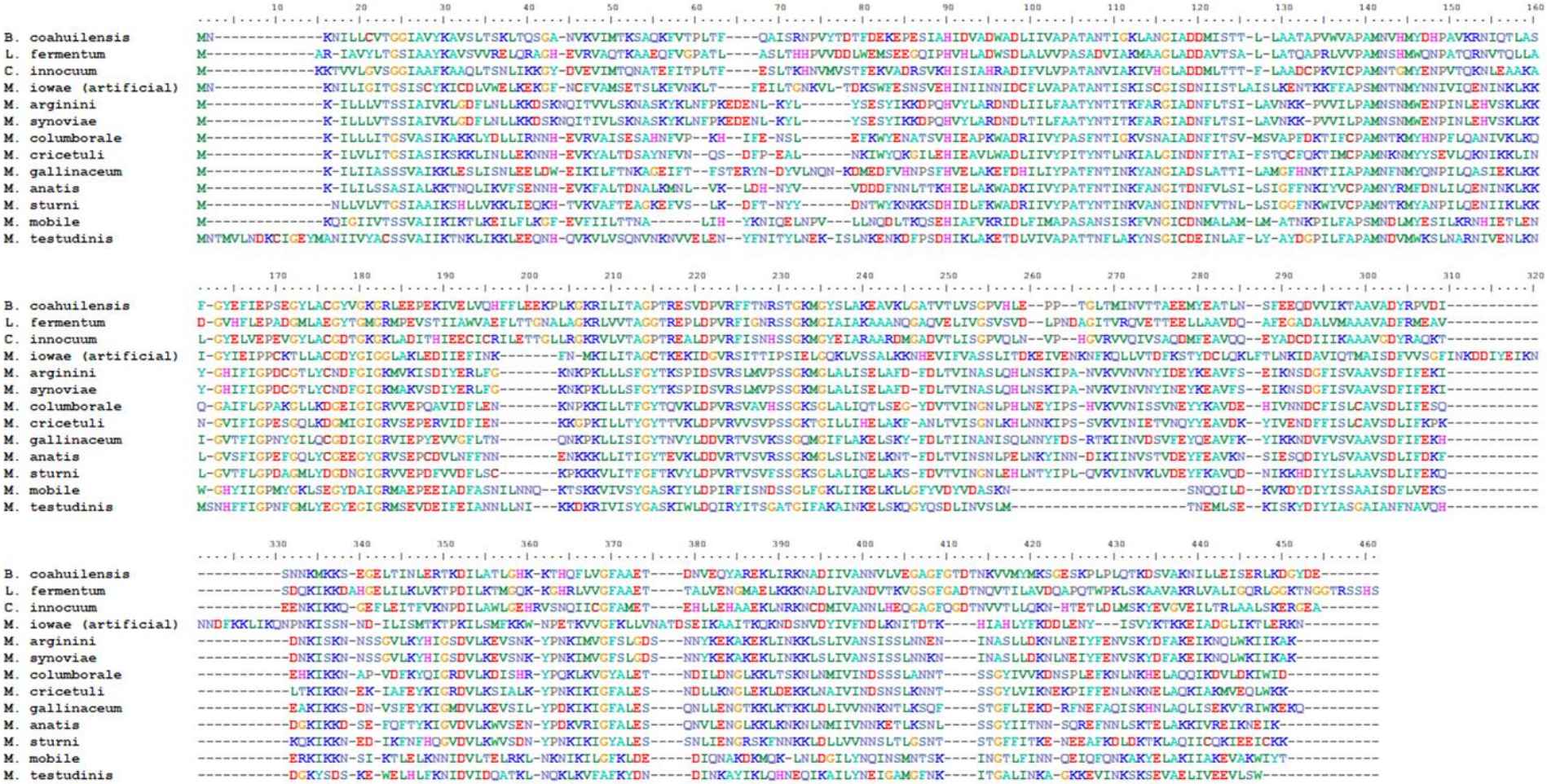
***Mycoplasma* species**	***p*-value**	**Motif locations**			
		Motif 1	**Motif 2**	**Motif 3**	**Motif 4**
M. sp. Ms02	5.40e-39	1-21	23-42	58-73	90-121
M. agalactiae	1.11e-57	1-21	23-43	58-73	90-121
M. alligatoris	1.89e-43	1-21	23-43	53-68	85-116
M. alvi	7.76e-20	2-22	23-43	58-73	-
M. anatis	6.14e-43	1-21	23-43	58-73	90-121
M. arginini	5.90e-52	1-21	23-43	57-72	89-120
M. bovigenitalium	6.45e-45	1-21	23-43	58-73	90-121
M. bovis	2.31e-56	1-21	23-43	58-73	90-121
*M. bovoculi**	5.64e-47	265-285	287-307	321-336	353-384
M. buteonis	2.98e-37	1-21	23-43	58-73	90-121
M. californicum	1.59e-43	1-21	23-43	58-73	90-121
M. canis	4.50e-39	1-21	23-43	57-72	89-120
M. capricolum subsp. *capricolum*	1.88e-24	6-26	30-50	63-78	-
M. capricolum subsp. *capripneumoniae*	5.67e-25	2-22	26-46	59-74	-
M. collis	1.28e-53	1-21	23-43	58-73	90-121
M. columbinum	4.89e-51	1-21	23-43	58-73	90-121
M. columborale	2.88e-38	1-21	23-43	57-72	89-120
*M. conjunctivae**	4.10e-51	266-286	288-308	322-337	354-385
M. cricetuli	8.29e-41	1-21	23-43	57-72	89-120
M. crocodyli	3.68e-47	1-21	23-43	58-73	90-121
*M. dispar**	2.73e-58	268-288	290-310	324-339	356-387
M. felifaucium	5.30e-54	1-21	23-43	58-73	90-121
M. felis	4.01e-40	3-23	25-45	60-75	92-123
M. fermentans	1.28e-41	1-21	23-43	58-73	90-121
*M. flocculare**	3.42e-55	267-287	289-309	323-338	355-386
M. gallinarum	1.59e-45	1-21	22-42	57-72	89-120
M. gallisepticum	1.35e-26	9-29	31-51	66-81	-
M. genitalium	2.60e-29	2-22	25-45	60-75	-
*M. hyopneumoniae**	3.07e-57	267-287	289-309	323-338	355-386
M. hyorhinis	4.62e-51	7-27	36-56	70-85	102-133
M. imitans	9.80e-25	9-29	31-51	66-81	-
M. iners	2.08e-53	1-21	23-43	58-73	90-121

***Mycoplasma* species**	***p*-value**	**Motif locations**			
		Motif 1	**Motif 2**	**Motif 3**	**Motif 4**
M. iowae	1.92e-23	11-31	33-53	68-83	-
M. leachii	3.03e-24	2-22	26-46	59-74	-
M. leonicaptivi	9.92e-43	1-21	23-43	58-73	90-121
M. lipofaciens	3.21e-47	1-21	23-43	58-73	90-121
M. mobile	7.43e-34	1-21	23-43	58-73	-
M. molare	1.02e-55	1-21	23-43	58-73	90-121
M. mycoides subsp. *capri*	8.30e-22	2-22	26-46	59-74	-
M. mycoides subsp. *mycoides*	1.38e-24	6-26	30-50	63-78	-
M. opalescens	7.78e-42	1-21	23-43	58-73	90-121
*M. ovipneumoniae**	4.21e-49	268-288	290-310	324-339	356-387
M. penetrans	2.96e-30	14-34	36-56	71-86	-
M. pirum	4.84e-25	7-27	28-48	63-78	-
M. pneumoniae	1.92e-27	2-22	25-45	133-148	-
M. primatum	2.97e-51	1-21	23-43	58-73	90-121
M. pulmonis	8.18e-55	1-21	23-43	58-73	90-121
M. putrefaciens	5.29e-22	2-22	26-46	60-75	-
M. simbae	2.38e-48	1-21	23-43	58-73	90-121
M. sturni	2.80e-43	1-21	23-43	58-73	90-121
M. synoviae	5.90e-52	1-21	23-43	57-72	89-120
M. testudinis	5.80e-27	2-22	24-44	59-74	-
M. yeatsii	2.57e-21	2-22	26-46	60-75	-

*HAD-DPCK proteins

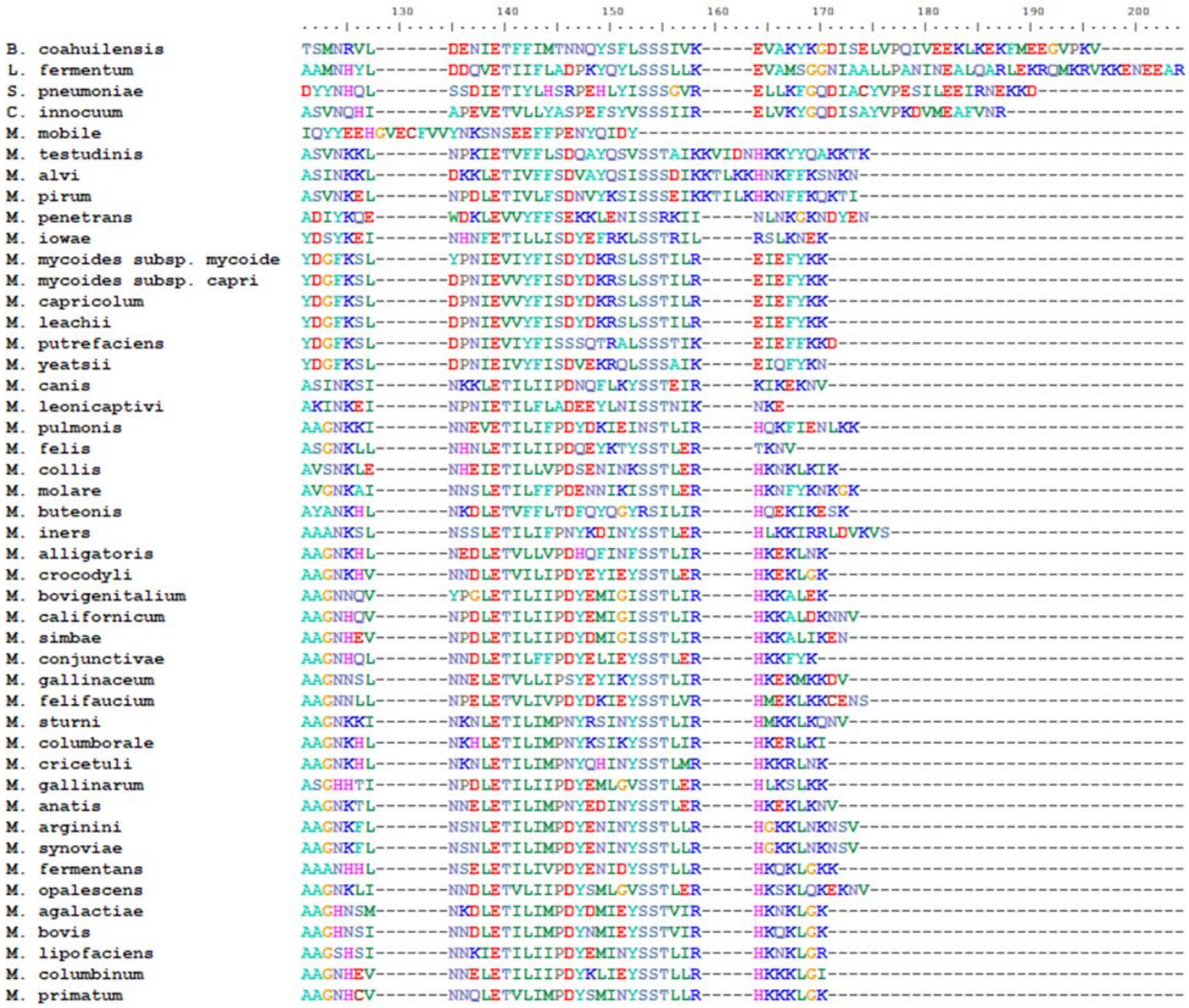
Supplementary Figure 1.1 The multiple sequence alignment (MSA) of the PanK type III amino acid sequences. Red box indicates deletions of *M. pulmonis*, *M. molare*, *M. mobile* and *M. testudinis* at position 236-238.

Supplementary Figure 1.2 The multiple sequence alignment (MSA) of the CoaBC amino acid sequences with *M. iowae* excluded. Red box indicates deletions of *M. mobile* and *M. testudinis* at position 256-275.



Supplementary Figure 1.3 The multiple sequence alignment (MSA) of the CoaBC amino acid sequences with the concatenated sequence of the PPCDC and PPCS sequences (artificial CoaBC sequence) of *M. iowae* included.

Supplementary Figure 1.4 The multiple sequence alignment (MSA) of the PPAT amino acid sequences.

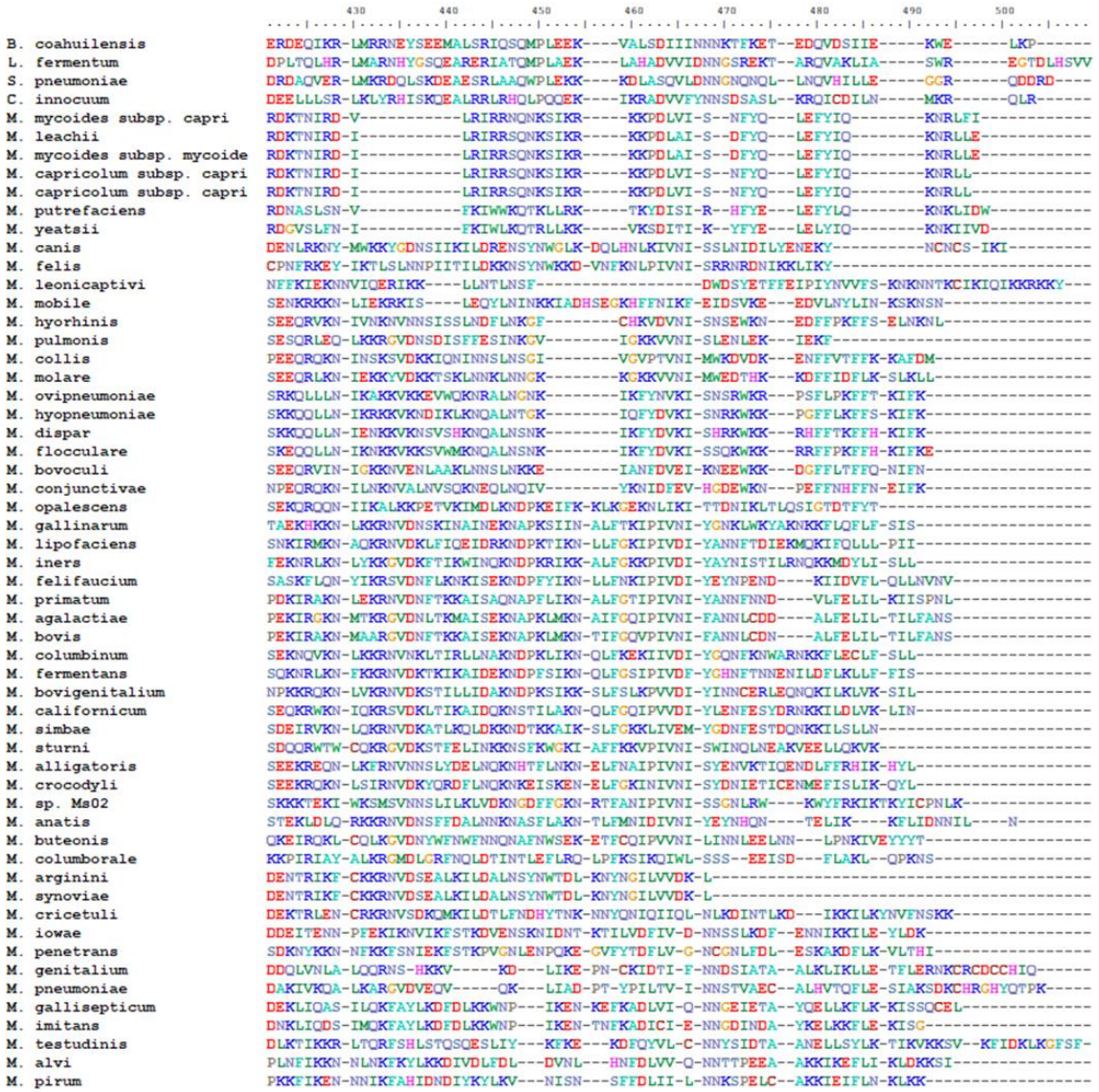


Supplementary Figure 1.4 continues The multiple sequence alignment (MSA) of the PPAT amino acid sequences.

Supplementary Figure 1.5 The multiple sequence alignment (MSA) of the DPCK and HAD-DPCK amino acid sequences.

Supplementary Figure 1.5 continues The multiple sequence alignment (MSA) of the DPCK and HAD-DPCK amino acid sequences.

Supplementary Figure 1.5 continues The multiple sequence alignment (MSA) of the DPCK and HAD-DPCK amino acid sequences.



Supplementary Figure 1.5 continues The multiple sequence alignment (MSA) of the DPCK and HAD-DPCK amino acid sequences.

Appendix 2 – Supplementary figures for Chapter 4

Supplementary Figure 2.1 The sequence alignment showing the sequencing results subsequent to the SDM of the *msDPCK* gene insert. The *msDPCK* gene sequence, along with the sequenced inserts of the isolated plasmids from colony 1 and 6 are shown, respectively. The red boxes indicate the TGA codons in the *msDPCK* gene.

Supplementary Figure 2.2 The sequence alignment showing the sequencing results subsequent to the sub-cloning of the SDM_*msDPCK* gene insert into a new pET28a(+) vector. The *msDPCK* gene and pET28a(+) vector sequences, along with the sequenced inserts of the isolated plasmids from colony 1, 2 and 3 are shown, respectively. The red boxes indicate the TGA codons in the *msDPCK* gene.